

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 9, 2004, 22:01:57 ; Search time 1121 Seconds
(without alignments)

11312.106 Million cell updates/sec

Title: US-10-054-841-3

Perfect score: 2985

Sequence: 1 gcgggcccgcagtggtgtg.....actgcggcccgcgcttgat 2985

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 337363 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29JUN04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	2985	100.0	2985	3	AAA09801 Human nuc
2	2983.4	99.9	2985	2	AAAX32266 Human nuc
3	2973	99.6	2987	2	AAA09802 Human nuc
4	2971.4	99.5	2987	2	AAAX32267 Human nuc
5	2901.8	97.2	5216	3	AAA40079 Human ERR
6	2901.8	97.2	5216	3	AAA40079 Human ERR
7	2826	88.0	5221	9	ADD69754 Human ERR
8	2497.8	83.7	3362	2	ADD69752 Human ERR
9	1559	52.2	1615	2	AAV47645 Steroid h
10	1418.2	47.5	1431	3	AAZ59741 Human oes
11	1377	46.1	1377	9	ABC23488 Human oes
12	1375.4	46.1	1377	9	AAZ59741 Human oes
13	719.6	24.1	1996	3	AAA40078 DNA encod
14	715	24.0	2807	2	AAA32265 Human ERR
15	715	24.0	2807	2	AAA32265 Human oes
16	713.4	23.9	2278	2	AAA09800 Human nuc
17	695	23.3	2153	1	AAAX9921 Human nuc
18	695	23.3	2153	1	AAAX9921 Human SBP
19	695	23.3	2153	6	ABZ34975 Sequence
20	443.2	14.8	536	6	ABZ34975 Human gen
21	428	14.3	622	3	AAV47646 Steroid h
22	427.6	14.3	622	2	AAA09804 Human nuc
23	424.4	14.2	2218	6	ABL64390 Nucleotid
					ABL64390 Stomach c

24	424.4	14.2	2218	6	ABK63889	Abx83889 Human cDN
25	421.4	14.1	2539	4	AA158567	AA158567 Human pol
26	421.4	14.1	2539	8	ADB48544	ADB48544 Novel hum
27	421.4	14.1	4250	6	ABZ34952	ABZ34952 Human nuc
28	401.6	13.5	2402	6	ABZ34954	ABZ34954 Human gen
29	401.6	13.5	2402	7	ABZ76223	ABZ76223 Human est
30	399.4	13.4	2249	1	AAAB0921	AAAB0921 Sequence
31	387.6	13.0	1305	7	ABZ68582	ABZ68582 Nucleotid
32	385.8	12.9	2559	6	AA160353	AA160353 Human pol
33	360.4	12.1	1599	6	AAV99427	AAV99427 Human NOV
34	356.2	11.9	9193	6	AAV46346	AAV46346 Human nuc
35	248.4	8.3	1131	6	AAV46346	AAV46346 Human nuc
36	204.6	6.9	403	2	AAAX32268	AAAX32268 Nucleotid
37	204.6	6.9	403	2	AAAX32268	AAAX32268 Human nuc
38	178.2	6.0	456	2	AAAX9922	AAAX9922 Human nuc
39	143.6	4.8	1455	7	AAV47602	AAV47602 Human SBP
40	143.6	4.8	1560	7	AAV47601	AAV47601 Drosophil
41	143.6	4.8	2529	4	ABL12869	ABL12869 Drosophil
42	141.4	4.7	1978	6	ABK89713	ABK89713 Oestrogen
43	136.6	4.6	1374	4	AAV29925	AAV29925 Human est
44	136.6	4.6	1380	4	AAV29925	AAV29925 Nucleotid
45	136.6	4.6	1770	9	ADB99351	ADB99351 Fusion pr

ALIGNMENTS

RESULT 1	AAA09801	standard; DNA; 2985 BP.
XX	XX	AAA09801;
AC	XX	05-JUL-2000 (first entry)
DT	XX	Human nuclear receptor nR2 gene.
DE	XX	Human nuclear receptor protein-2; nR2; physiological function; ds;
KW	XX	cell development and differentiation controller; gene expression.
OS	XX	Homo sapiens.
XX	XX	US6054295-A.
PN	XX	25-APR-2000.
PD	XX	26-AUG-1998; 98US-00141000.
PF	XX	27-AUG-1997; 97US-0057090P.
PR	XX	21-OCT-1997; 97US-0062922P.
PR	XX	19-MAR-1998; 98US-0078633P.
PA	XX	(MERI) MERCK & CO INC.
PI	XX	Chen F;
XX	XX	WPI, 2000-328352/28.
DR	XX	P-PSDB; AAY83823.
PT	XX	New polynucleotide encoding human nuclear receptor protein-1 is useful
PT	XX	for screening compounds that acts as modulators of cell differentiation,
XX	XX	cell development and physiological function.
XX	XX	Disclosure; Fig 5; 58pp; English.
XX	XX	This sequence represents the DNA encoding human nuclear receptor protein-
CC	XX	2 (nR2). The protein is a cell development and differentiation
CC	XX	controller which is useful for screening compounds that act as modulators
CC	XX	of cell differentiation, cell development and physiological functions or
CC	XX	for treating and diagnosing disorders associated with altered expression
XX	XX	of the DNA
XX	XX	Sequence 2985 BP; 794 A; 703 C; 605 G; 883 T; 0 U; 0 Other;
XX	XX	SD

Query Match 100.0%; Score 2985; DB 3; Length 2985;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGGGCCGCAAGTGTGTGAATTCGCTTGTCTACATGAGAAACATTGGTTAATTGCA 60
 1 GCGGGCCGCAAGTGTGTGAATTCGCTTGTCTACATGAGAAACATTGGTTAATTGCA 60
 61 CTGTGCTCTGTCAAGAACTTTGATTTATAGCTGGGGGTSCAAATAATGATGGCGGT 120
 61 CTGTGCTCTGTCAAGAACTTTGATTTATAGCTGGGGGTSCAAATAATGATGGCGGT 120
 121 CGCAGATGATTCGGTGAACCTTTGCTTCCTGATCTTTTCCCTGCTACAGAGAA 180
 121 CGCAGATGATTCGGTGAACCTTTGCTTCCTGATCTTTTCCCTGCTACAGAGAA 180
 181 AGCTTCTCTGAGAAATGTCAAAACAAAGATGACACATTGATTCGAGCTGTCCCTTCA 240
 181 AGCTTCTCTGAGAAATGTCAAAACAAAGATGACACATTGATTCGAGCTGTCCCTTCA 240
 241 TCAAGAGGAAACCTTCAGAGCCAGCTCCCTGACGAGAGGCTGACCAAGCCCTG 300
 241 TCAAGAGGAAACCTTCAGAGCCAGCTCCCTGACGAGAGGCTGACCAAGCCCTG 300
 301 GTGGCTCTTCAGAGCCAGCTGAGGCTTACAGTTCAACCATGATGAGCCATGAGAC 360
 301 GTGGCTCTTCAGAGCCAGCTGAGGCTTACAGTTCAACCATGATGAGCCATGAGAC 360
 361 TTGACTGCGCACTCTTCACTTCTGTCTCTATTCCTGGAGGTAGTGGGCTGTGACGA 420
 361 TTGACTGCGCACTCTTCACTTCTGTCTCTATTCCTGGAGGTAGTGGGCTGTGACGA 420
 421 AACTGTATGATGATCTGCTTCAGACACATTTGTAAGATCCCAAGCAAGGTGTAATCA 480
 421 AACTGTATGATGATCTGCTTCAGACACATTTGTAAGATCCCAAGCAAGGTGTAATCA 480
 481 TGCTCAACTGATGATCCCAAGAGCTGTGTGTAGTGTGTGTGATCGCTTCTGGGTACC 540
 481 TGCTCAACTGATGATCCCAAGAGCTGTGTGTAGTGTGTGTGATCGCTTCTGGGTACC 540
 541 ACTATGGGGTATGATCATGTGAAGCTGTGAAGGCATTTCTCAAGAGCAATTCAGAGCA 600
 541 ACTATGGGGTATGATCATGTGAAGCTGTGAAGGCATTTCTCAAGAGCAATTCAGAGCA 600
 601 AATATGAAATACAGCTGCTCCGACCAAGATGATGTAATACAAACGCGAGACGTAAAT 660
 601 AATATGAAATACAGCTGCTCCGACCAAGATGATGTAATACAAACGCGAGACGTAAAT 660
 661 CCTGCCAGGCTTCCGCTTCATGAAAGTGTAAAGTGGGCAATGCTGAAAGAAAGGGGTGC 720
 661 CCTGCCAGGCTTCCGCTTCATGAAAGTGTAAAGTGGGCAATGCTGAAAGAAAGGGGTGC 720
 721 GTCTTGAAGAGTATGCTGAGAGTCCGAGAAATCAAGCCAGAGATGATGCGAGAAACA 780
 721 GTCTTGAAGAGTATGCTGAGAGTCCGAGAAATCAAGCCAGAGATGATGCGAGAAACA 780
 781 GCCCATACCTGAACCTCAGCTGCTTCAGCCAGCAAAAGCCATATPACAAAGTTGTCT 840
 781 GCCCATACCTGAACCTCAGCTGCTTCAGCCAGCAAAAGCCATATPACAAAGTTGTCT 840
 841 CACATTTGTGTGTGCTGAACCGGAGAAATCTATGTCATGCTGAGCCCTTACGTCCCG 900
 841 CACATTTGTGTGTGCTGAACCGGAGAAATCTATGTCATGCTGAGCCCTTACGTCCCG 900
 901 ACAGTGACATTAAGCCCTCACTACATGCTGATGCTTGGCCGACGAGAGTTGTGTGTTA 960
 901 ACAGTGACATTAAGCCCTCACTACATGCTGATGCTTGGCCGACGAGAGTTGTGTGTTA 960
 961 TCATTTGATGGGCGAAGCATATTCAGAGCTTCTCCAGCTGTCTGCGGCGAGCATGGA 1020
 961 TCATTTGATGGGCGAAGCATATTCAGAGCTTCTCCAGCTGTCTGCGGCGAGCATGGA 1020

1021 GCCTTCGAGAGTGTGTGAATTTGATCTGTGCTGCTGATACGGGTCTCTT 1080
 1021 GCCTTCGAGAGTGTGTGAATTTGATCTGTGCTGCTGATACGGGTCTCTT 1080
 1081 CATTGAGATGATCTGTCTATGACAGCATATPATAATGACGAACAGTCCAAAT 1140
 1081 CATTGAGATGATCTGTCTATGACAGCATATPATAATGACGAACAGTCCAAAT 1140
 1141 TGCAGAGCTCTTGTGATTAATATGCTATCTGTGAGCTGTGTAAGAAATACAGAGCA 1200
 1141 TGCAGAGCTCTTGTGATTAATATGCTATCTGTGAGCTGTGTAAGAAATACAGAGCA 1200
 1201 TGAAGCTGAAAAAAGAAATTTGTCACTCCCTGAAAGCTATAGCTCTTGTATCAACT 1260
 1201 TGAAGCTGAAAAAAGAAATTTGTCACTCCCTGAAAGCTATAGCTCTTGTATCAACT 1260
 1261 CCATGCAATGAAAGATGTTGAAAGCCCTTCAGAAAGCTTCAGAGATGCTTACATGAGCGC 1320
 1261 CCATGCAATGAAAGATGTTGAAAGCCCTTCAGAAAGCTTCAGAGATGCTTACATGAGCGC 1320
 1321 TGCAGATTAATGAAGCTGCGCAGCATGGAAGACCTCTGAGAGCTGGGAGAGTGTGA 1380
 1321 TGCAGATTAATGAAGCTGCGCAGCATGGAAGACCTCTGAGAGCTGGGAGAGTGTGA 1380
 1381 TGACACTGCCACTCTTATGAGAGACCTCTACAAAGCCGTGAGCTGCAAGATGCTGA 1440
 1381 TGACACTGCCACTCTTATGAGAGACCTCTACAAAGCCGTGAGCTGCAAGATGCTGA 1440
 1441 AACTGAGAGGAAAGTCCCAATGCAAAACCTTTTGGAAATGTTGAGGCGAAAGTCT 1500
 1441 AACTGAGAGGAAAGTCCCAATGCAAAACCTTTTGGAAATGTTGAGGCGAAAGTCT 1500
 1501 GACTTAAAGCTCCCTGGGCTCTCCATCTTCAATGTTGAAAAAGGAAATTAACCCAG 1560
 1501 GACTTAAAGCTCCCTGGGCTCTCCATCTTCAATGTTGAAAAAGGAAATTAACCCAG 1560
 1561 AGTATGTCGAGAAACCTTAGGTTAGTTAACAAATCAAAATCAACAGACTGACTG 1620
 1561 AGTATGTCGAGAAACCTTAGGTTAGTTAACAAATCAAAATCAACAGACTGACTG 1620
 1621 AATAATTTGACAGACATGATGAGAGCTTCAGATTCCTGCAATGAGTCTGATGAGT 1680
 1621 AATAATTTGACAGACATGATGAGAGCTTCAGATTCCTGCAATGAGTCTGATGAGT 1680
 1681 TCTTCTACTTCTCCATCATCTTCTTCTCTTCTTCTTCCACATTTCTTCTCTTA 1740
 1681 TCTTCTACTTCTCCATCATCTTCTTCTCTTCTTCTTCCACATTTCTTCTCTTA 1740
 1741 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
 1741 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
 1801 CATTCCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 1801 CATTCCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 1861 TCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
 1861 TCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
 1921 AAAAATCCTCCCTCCCTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
 1921 AAAAATCCTCCCTCCCTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
 1981 TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2040
 1981 TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2040
 2041 CTTTAAAAAGAGTCTCTTAACTGAGAGATGGAAGCCAGCCGCAAGAGATGAGGA 2100
 2041 CTTTAAAAAGAGTCTCTTAACTGAGAGATGGAAGCCAGCCGCAAGAGATGAGGA 2100
 2101 TCATTAATATGATGCGAGTGAACCTTATGTGAAACCATACGTTCCCAATGACTAAGGA 2160

```

Db      2101 TCCATATATGATGATGAGCCAGTGAACCTTATGTAACCAATCCGCCCCAATGATAGTAAGAA 2160
QY      2161 TCAAGAGAGAGAACCAAGCTTCTTAAGTACAGTGCAACATATACAAATTAAGTACAGT 2220
Db      2161 TCAAGAGAGAGAACCAAGCTTCTTAAGTACAGTGCAACATATACAAATTAAGTACAGT 2220
QY      2221 GCGATTTGATTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
Db      2221 GCGATTTGATTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY      2281 CTGCTCTTATCATCTGCTTTTTCATCATATCATATCATATCATATCATATCATATCATAT 2340
Db      2281 CTGCTCTTATCATCTGCTTTTTCATCATATCATATCATATCATATCATATCATATCATAT 2340
QY      2341 TTTTTCAGTCTTCCAGTATTTGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 2400
Db      2341 TTTTTCAGTCTTCCAGTATTTGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 2400
QY      2401 TTTTTCAGTCTTCCAGTATTTGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 2460
Db      2401 TTTTTCAGTCTTCCAGTATTTGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 2460
QY      2461 GGCCTTTGAACAATTAAGCAATATATATATATATATATATATATATATATATATATATAT 2520
Db      2461 GGCCTTTGAACAATTAAGCAATATATATATATATATATATATATATATATATATATATAT 2520
QY      2521 AAATGTTTGTGCTTCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Db      2521 AAATGTTTGTGCTTCTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
QY      2581 CCATTTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2640
Db      2581 CCATTTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2640
QY      2641 ATACAGGATGCAATTTATATATATATATATATATATATATATATATATATATATATAT 2700
Db      2641 ATACAGGATGCAATTTATATATATATATATATATATATATATATATATATATATATAT 2700
QY      2701 AGCTCAGTTTACCTAAATGTTTACCTAGCAACAAAGATGAGATTAACCTACAGTACATC 2760
Db      2701 AGCTCAGTTTACCTAAATGTTTACCTAGCAACAAAGATGAGATTAACCTACAGTACATC 2760
QY      2761 TGAATGTCATATATATATATATATATATATATATATATATATATATATATATATATATAT 2820
Db      2761 TGAATGTCATATATATATATATATATATATATATATATATATATATATATATATATATAT 2820
QY      2821 GCACTCTCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Db      2821 GCACTCTCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
QY      2881 AATGATATATCTGCTGTTGCAAAAGTCAAGAGCTTGGCTGAGAGAGAGAGAGAGAGAGAG 2940
Db      2881 AATGATATATCTGCTGTTGCAAAAGTCAAGAGCTTGGCTGAGAGAGAGAGAGAGAGAGAG 2940
QY      2941 TATTTATTTAAGAGATTTCCACACACTGCGCGCCGCTTGAT 2985
Db      2941 TATTTATTTAAGAGATTTCCACACACTGCGCGCCGCTTGAT 2985

```

```

RESULT 2
AAK32266
ID      AAK32266 standard; DNA; 2985 BP.
XX
XX      AAK32266;
XX
XX      15-JUN-1999 (first entry)
XX
XX      Human nuclear receptor protein nNR2 encoding DNA.
XX      Nuclear receptor; nNR1; nNR2; cell differentiation; human; ss.
XX
XX      Homo sapiens.

```

```

XX      WO9910367-A1.
XX
XX      04-MAR-1999.
XX
XX      27-AUG-1998; 98WO-US017826.
XX
XX      27-AUG-1997; 97US-0057090P.
XX      21-OCT-1997; 97US-0062902P.
XX      19-MAR-1998; 98US-0078633P.
XX
XX      (MERI ) MERCK & CO INC.
XX
XX      Chen F;
XX
XX      WPI: 1999-190586/16.
XX      P-PSDB; AAK03837.
XX
XX      Newly purified DNA polynucleotides encoding human nuclear trans-acting
XX      receptor proteins - useful in the diagnosis, treatment and prophylaxis of
XX      cell differentiation, development and physiological function.
XX
XX      Claim 31; Page 48-49; 82pp; English.
XX
XX      The invention relates to DNA molecules encoding human nuclear receptor
XX      (nNR) proteins nNR1 and nNR2. The nNR proteins (including mutants and/or
XX      fragments) form pharmaceutical compositions that are useful in the
XX      diagnosis, treatment and prophylaxis of cell differentiation, development
XX      and physiological function. The proteins are also useful for identifying
XX      downstream target genes and ligands regulating their activity. In
XX      particular, fusion constructs (especially glutathione S-transferase (GST)
XX      -nNR1 and/or GST-nNR2) expressing fusion proteins are useful in screening
XX      for (ant)agonists that are useful as modulators in cell differentiation,
XX      development and physiological function. Antibodies for the nNR proteins
XX      are useful for measuring the levels of these proteins. The present
XX      sequence represents a DNA encoding a human nNR2 protein
XX
XX      Sequence 2985 BP; 794 A; 702 C; 606 G; 893 T; 0 U; 0 Other;
XX
XX      Query Match 99.9%; Score 2983.4; DB 2; Length 2985;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 2984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```


Db 961 TCATTGATGGGCGAGACATATTCAGGCTTCTCCACGCTGTCCCTGGCCGACCAAGTGA 1020
QY 1021 GCCCTTGCAGAGAGCTTGGATGGAAATTTTGATCTTGGTGTGTATACCGGCTCTTT 1080
Db 1021 GCCCTTGCAGAGAGCTTGGATGGAAATTTTGATCTTGGTGTGTATACCGGCTCTTT 1080
QY 1081 CATTGAGAGATGAACCTGTCTATGACAGCATATATATATGAGAGAGAACCAAGTCCAAAT 1140
Db 1081 CATTGAGAGATGAACCTGTCTATGACAGCATATATATATGAGAGAGAACCAAGTCCAAAT 1140
QY 1141 TAGAGAGCTCTTGTATTAATATATGCTATCCGACAGCTGTAAAGAAATACAGAGCA 1200
Db 1141 TAGAGAGCTCTTGTATTAATATATGCTATCCGACAGCTGTAAAGAAATACAGAGCA 1200
QY 1201 TGAAGCTGGAAAAAGAGAAATTTGTACCTCTCAAGGCTATAGCTTTGCTATTCAGACT 1260
Db 1201 TGAAGCTGGAAAAAGAGAAATTTGTACCTCTCAAGGCTATAGCTTTGCTATTCAGACT 1260
QY 1261 CCATGCAATAGAGATGTGTGAAGCCCTTCAGAGAGCTCAGGATGTCTACATGAGAGGC 1320
Db 1261 CCATGCAATAGAGATGTGTGAAGCCCTTCAGAGAGCTCAGGATGTCTACATGAGAGGC 1320
QY 1261 CCATGCAATAGAGATGTGTGAAGCCCTTCAGAGAGCTCAGGATGTCTACATGAGAGGC 1320
Db 1261 CCATGCAATAGAGATGTGTGAAGCCCTTCAGAGAGCTCAGGATGTCTACATGAGAGGC 1320
QY 1321 TGACGATTAATGAAAGCTGAGCCAGCATG--GAAGACCTCGTCGAGCTGGCAGAGTCT 1378
Db 1321 TGACGATTAATGAAAGCTGAGCCAGCATG--GAAGACCTCGTCGAGCTGGCAGAGTCT 1378
QY 1379 GATGACACTGCGACTCTGAGGAGACCTCTACCAAGCCGTGCGACATTTCTACAAAT 1438
Db 1379 GATGACACTGCGACTCTGAGGAGACCTCTACCAAGCCGTGCGACATTTCTACAAAT 1438
QY 1381 GATGACACTGCGACTCTGAGGAGACCTCTACCAAGCCGTGCGACATTTCTACAAAT 1440
Db 1381 GATGACACTGCGACTCTGAGGAGACCTCTACCAAGCCGTGCGACATTTCTACAAAT 1440
QY 1439 CAAATAGAGAGCAAAAGTCCCAATGCAAACTTTTGGAAATGTTGAGAGCCAAAGT 1498
Db 1439 CAAATAGAGAGCAAAAGTCCCAATGCAAACTTTTGGAAATGTTGAGAGCCAAAGT 1498
QY 1441 CAAATAGAGAGCAAAAGTCCCAATGCAAACTTTTGGAAATGTTGAGAGCCAAAGT 1500
Db 1441 CAAATAGAGAGCAAAAGTCCCAATGCAAACTTTTGGAAATGTTGAGAGCCAAAGT 1500
QY 1499 CTGACTAAAAAGCTCCCTGGGCTTCCCATCTTCAATGTTGAAAAAGGAAAAATAAACCA 1558
Db 1499 CTGACTAAAAAGCTCCCTGGGCTTCCCATCTTCAATGTTGAAAAAGGAAAAATAAACCA 1558
QY 1501 CTGACTAAAAAGCTCCCTGGGCTTCCCATCTTCAATGTTGAAAAAGGAAAAATAAACCA 1560
Db 1501 CTGACTAAAAAGCTCCCTGGGCTTCCCATCTTCAATGTTGAAAAAGGAAAAATAAACCA 1560
QY 1559 AGAGTGATGTCAGAGAACTTAGAGTTTAGTTAACAACTCAAAATCAAGACCTGGAC 1618
Db 1559 AGAGTGATGTCAGAGAACTTAGAGTTTAGTTAACAACTCAAAATCAAGACCTGGAC 1618
QY 1561 AGAGTGATGTCAGAGAACTTAGAGTTTAGTTAACAACTCAAAATCAAGACCTGGAC 1620
Db 1561 AGAGTGATGTCAGAGAACTTAGAGTTTAGTTAACAACTCAAAATCAAGACCTGGAC 1620
QY 1619 TGATTAATTTAGAGAGAGCACTAGAGAGCTTTCAGATTCGCCATAGATTCCTGATGA 1678
Db 1619 TGATTAATTTAGAGAGAGCACTAGAGAGCTTTCAGATTCGCCATAGATTCCTGATGA 1678
QY 1621 TGATTAATTTAGAGAGAGCACTAGAGAGCTTTCAGATTCGCCATAGATTCCTGATGA 1680
Db 1621 TGATTAATTTAGAGAGAGCACTAGAGAGCTTTCAGATTCGCCATAGATTCCTGATGA 1680
QY 1679 GTTCTTCTACTTTCCTCAATCTTCTTCTCTTCTTCTTCCCAATTTCTCTTCTCTT 1738
Db 1679 GTTCTTCTACTTTCCTCAATCTTCTTCTCTTCTTCTTCCCAATTTCTCTTCTCTT 1738
QY 1681 GTTCTTCTACTTTCCTCAATCTTCTTCTCTTCTTCTTCCCAATTTCTCTTCTCTT 1740
Db 1681 GTTCTTCTACTTTCCTCAATCTTCTTCTCTTCTTCTTCCCAATTTCTCTTCTCTT 1740
QY 1739 TATTTTCTCTCTTCTTCTTCTTCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1798
Db 1739 TATTTTCTCTCTTCTTCTTCTTCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1798
QY 1741 TATTTTCTCTCTTCTTCTTCTTCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
Db 1741 TATTTTCTCTCTTCTTCTTCTTCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
QY 1799 CCGATCTCTCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1858
Db 1799 CCGATCTCTCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1858
QY 1801 CCGATCTCTCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
Db 1801 CCGATCTCTCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
QY 1859 ATTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1918
Db 1859 ATTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1918
QY 1861 ATTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1920
Db 1861 ATTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1920
QY 1919 AAAAAAATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1978
Db 1919 AAAAAAATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1978
QY 1921 AAAAAAATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
Db 1921 AAAAAAATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
QY 1979 CTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2038
Db 1979 CTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2038
QY 1981 CTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2040
Db 1981 CTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2040
QY 2039 AACTTTTAAAGAGTCTCTTAACTGAAAGAGATGAAAGCCAGCCCTGCAAAAGATGGA 2098
Db 2039 AACTTTTAAAGAGTCTCTTAACTGAAAGAGATGAAAGCCAGCCCTGCAAAAGATGGA 2098
QY 2041 AACTTTTAAAGAGTCTCTTAACTGAAAGAGATGAAAGCCAGCCCTGCAAAAGATGGA 2100
Db 2041 AACTTTTAAAGAGTCTCTTAACTGAAAGAGATGAAAGCCAGCCCTGCAAAAGATGGA 2100

QY 2099 GATCCATTAATATGAGATGCCAGTGAATTAATTTGAAACCATACGTCCTCCCATATGACTAAG 2158
Db 2101 GATCCATTAATATGAGATGCCAGTGAATTAATTTGAAACCATACGTCCTCCCATATGACTAAG 2160
QY 2159 AATCAAGAGAGAGAACCAAGTCTCTTAAAGTACAGTGAACATATACAAATTTGACTGA 2218
Db 2161 AATCAAGAGAGAGAACCAAGTCTCTTAAAGTACAGTGAACATATACAAATTTGACTGA 2220
QY 2219 GTGCAATTAATGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2278
Db 2221 GTGCAATTAATGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY 2279 GGTGCTCTCTTATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2338
Db 2281 GGTGCTCTCTTATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2340
QY 2339 GTTTTTCAGTCTTCCAGGATTTTGTATTTGATTTGATTTGATTTGATTTGATTTGAT 2398
Db 2341 GTTTTTCAGTCTTCCAGGATTTTGTATTTGATTTGATTTGATTTGATTTGATTTGAT 2400
QY 2399 AGTTTAAGCTTTATTTCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2458
Db 2401 AGTTTAAGCTTTATTTCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2460
QY 2459 CTGGCTTTGAAACAATTTAGCAACAATTTAGCAACAATTTAGCAACAATTTAGCAACAATTT 2518
Db 2461 CTGGCTTTGAAACAATTTAGCAACAATTTAGCAACAATTTAGCAACAATTTAGCAACAATTT 2520
QY 2519 AAAAAATGTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2578
Db 2521 AAAAAATGTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2580
QY 2579 AGCCATTAATTAATTTAGGAAATTTAGTGAATTTAGTGAATTTAGTGAATTTAGTGAATTT 2638
Db 2581 AGCCATTAATTAATTTAGGAAATTTAGTGAATTTAGTGAATTTAGTGAATTTAGTGAATTT 2640
QY 2639 TGAGCTCCAGTTTACCTTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2758
Db 2641 TGAGCTCCAGTTTACCTTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2760
QY 2701 TGAGCTCCAGTTTACCTTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2818
Db 2701 TGAGCTCCAGTTTACCTTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2820
QY 2759 TCGAGTATGATCAATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2878
Db 2761 TCGAGTATGATCAATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
QY 2819 GTGCACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2938
Db 2821 GTGCACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2940
QY 2879 GGAATGATGATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2980
Db 2881 GGAATGATGATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2982
QY 2939 AGTATTCATTTAGAGATTTCCACCACTGGGAGCCCGGCTTGAAT 2985
Db 2941 AGTATTCATTTAGAGATTTCCACCACTGGGAGCCCGGCTTGAAT 2987
Db 2941 AGTATTCATTTAGAGATTTCCACCACTGGGAGCCCGGCTTGAAT 2987

RESULT 4

AAAX32267

ID AAX32267 standard; DNA; 2987 BP.

AC AAX32267;

DT 15-JUN-1999 (first entry)

DE Human nuclear receptor protein NR2-1 encoding DNA.

KW Nucleic receptor; NR1; NR2; cell differentiation; human; ss.

XX Homo sapiens.
 OS MO9910367-A1.
 XX 04-MAR-1999.
 XX PD
 XX PF 27-AUG-1998; 98WO-US017826.
 XX PR 27-AUG-1997; 97US-0057090P.
 XX PR 21-OCT-1997; 97US-0062902P.
 XX PR 19-MAR-1998; 98US-0078633P.
 XX PA (MERI) MERCK & CO INC.
 XX PI Chen F;
 XX DR WPI, 1999-190586/16.
 XX DR P-PSDB; AAY03838.
 XX PT Newly purified DNA polynucleotides encoding human nuclear trans-acting
 PT receptor proteins - useful in the diagnosis, treatment and prophylaxis of
 PT cell differentiation, development and physiological function.
 XX PS Disclosure; Fig 7A-C; 82pp; English.
 XX CC The invention relates to DNA molecules encoding human nuclear receptor
 CC (nNR) proteins nNR1 and nNR2. The nNR proteins (including mutants and/or
 CC fragments) form pharmaceutical compositions that are useful in the
 CC diagnosis, treatment and prophylaxis of cell differentiation, development
 CC and physiological function. The proteins are also useful for identifying
 CC downstream target genes and ligands regulating their activity. In
 CC particular, fusion constructs (especially glutathione S-transferase (GST)
 CC -nNR1 and/or GST-nNR2) expressing fusion proteins are useful in screening
 CC for (anti)agonists that are useful as modulators in cell differentiation,
 CC development and physiological function. Antibodies for the nNR proteins
 CC are useful for measuring the levels of these proteins. The present
 CC sequence represents a DNA encoding a human nNR2-1 protein
 XX
 SQ Sequence 2987 BP; 795 A; 702 C; 607 G; 883 T; 0 U; 0 Other;

Query Match 99.5%; Score 2971.4; DB 2; Length 2987;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2984; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
 QY 1 GCGGGCCGCCAGTGTGGATTCGGCTTGTCACTAGAGAACATTTGTTAATTCGA 60
 DB 1 GCGGGCCGCCAGTGTGGATTCGGCTTGTCACTAGAGAACATTTGTTAATTCGA 60
 QY 61 CTGGCTCTGTCAAGAACTTGAATTTATAGCTGGGGTGCACAAATATGCTTGGCGGT 120
 DB 61 CTGGCTCTGTCAAGAACTTGAATTTATAGCTGGGGTGCACAAATATGCTTGGCGGT 120
 QY 121 CGACATGATTCGGTGAACCTTTCCTTCCTGATCTTTTCCCTGCATCAGAGAG 180
 DB 121 CGACATGATTCGGTGAACCTTTCCTTCCTGATCTTTTCCCTGCATCAGAGAG 180
 QY 181 AGCTTCTCTGCAAGATGTCAAGAAAGATGACACATTGATTCAGCTGTTCCTTCA 240
 DB 181 AGCTTCTCTGCAAGATGTCAAGAAAGATGACACATTGATTCAGCTGTTCCTTCA 240
 QY 241 TCAAGACGGAACCTTCCAGCCAGCCCTCCTGACGACAGCGTCAACACCAACAGCCCTG 300
 DB 241 TCAAGACGGAACCTTCCAGCCAGCCCTCCTGACGACAGCGTCAACACCAACAGCCCTG 300
 QY 301 GTGGCTCTTTCAGACGCGCAGTGGAGGTACAGTTCAACATGAATGGCCATCAGAACGAG 360
 DB 301 GTGGCTCTTTCAGACGCGCAGTGGAGGTACAGTTCAACATGAATGGCCATCAGAACGAG 360
 QY 361 TTGACTTCGCACTCTCTTACCCCTTTCCTCTTACTCTGGAGGTAGTGGGCTGTGACGA 420
 DB 361 TTGACTTCGCACTCTCTTACCCCTTTCCTCTTACTCTGGAGGTAGTGGGCTGTGACGA 420

QY 421 AACTGTATGATGACTGCTCCAGACCAATGTTGAAGATCCCGACCAAGTGTGAATCA 480
 DB 421 AACTGTATGATGACTGCTCCAGACCAATGTTGAAGATCCCGACCAAGTGTGAATCA 480
 QY 481 TGTCTCACTGATGCGCCCAAGAGACTGTGTAGTGTGGTGAATTCGCTTCTGGTACC 540
 DB 481 TGTCTCACTGATGCGCCCAAGAGACTGTGTAGTGTGGTGAATTCGCTTCTGGTACC 540
 QY 541 ACTATGGGGTATGATCATGTGAAGGCTGCAAGGCAATTCGCAAGGCAATTCGAAGCA 600
 DB 541 ACTATGGGGTATGATCATGTGAAGGCTGCAAGGCAATTCGCAAGGCAATTCGAAGCA 600
 QY 601 ATATGAATACAGTCCCTCCAGCAAGATGATGAATATCAAAAGGCGAGACTTAAT 660
 DB 601 ATATGAATACAGTCCCTCCAGCAAGATGATGAATATCAAAAGGCGAGACTTAAT 660
 QY 661 CCGCCAGGCTTGGCGCTTCATGAAGTATTAAGTGGGCAATGCTGAAGAGGGGTGC 720
 DB 661 CCGCCAGGCTTGGCGCTTCATGAAGTATTAAGTGGGCAATGCTGAAGAGGGGTGC 720
 QY 721 GTCTTGAACAGATGACGTGAGAGGTGCGGAGAGTCAAGCGAGATAGTGGGAGACA 780
 DB 721 GTCTTGAACAGATGACGTGAGAGGTGCGGAGAGTCAAGCGAGATAGTGGGAGACA 780
 QY 781 GCCCATACCTGAACCTTCAAGCTGTGACGCCAGCCAAAGCCATATTAACAAGATTGCT 840
 DB 781 GCCCATACCTGAACCTTCAAGCTGTGACGCCAGCCAAAGCCATATTAACAAGATTGCT 840
 QY 841 CAGATTTGTTGGTGGCGTGAACCCGGAAGATCTATGCAATGCTGACCTTCTGTCCCG 900
 DB 841 CAGATTTGTTGGTGGCGTGAACCCGGAAGATCTATGCAATGCTGACCTTCTGTCCCG 900
 QY 901 ACACTGACATCAAAAGCCCTCACTACCTGTGTGACTTGGCGGACCGAGTGTGGTGA 960
 DB 901 ACACTGACATCAAAAGCCCTCACTACCTGTGTGACTTGGCGGACCGAGTGTGGTGA 960
 QY 961 TCATTTGATGGGCGAAGCATATTCAGGCTTCTCAAGCTGTCCCTGGCGGACCAAGATGA 1020
 DB 961 TCATTTGATGGGCGAAGCATATTCAGGCTTCTCAAGCTGTCCCTGGCGGACCAAGATGA 1020
 QY 1021 GCGTTCGCAAGAGTCTTGTGATGGAATTTTGTATCCTTGTGTGTATCCGGTCTCTT 1080
 DB 1021 GCGTTCGCAAGAGTCTTGTGATGGAATTTTGTATCCTTGTGTGTATCCGGTCTCTT 1080
 QY 1081 CATTGAGATGATGATCTGTATGACAGACATTAATATGAAGCAAGCCAGTCCAAAT 1140
 DB 1081 CATTGAGATGATGATCTGTATGACAGACATTAATATGAAGCAAGCCAGTCCAAAT 1140
 QY 1141 TAGAGGCTCTTCTTGAATCTAATATATGCTATCTGACGCTGTGAAGAAATACAGAGCA 1200
 DB 1141 TAGAGGCTCTTCTTGAATCTAATATATGCTATCTGACGCTGTGAAGAAATACAGAGCA 1200
 QY 1201 TGAAGCTGGAAGAAAGAAATTTGTCAACCTTCAAGCTATAGCTCTTCAATTCAGACT 1260
 DB 1201 TGAAGCTGGAAGAAAGAAATTTGTCAACCTTCAAGCTATAGCTCTTCAATTCAGACT 1260
 QY 1261 CCATGCACTAAGAAAGATGTTGAAGCGCTTCAAGAGCTTCAAGAGTCTTCAAGAGCGC 1320
 DB 1261 CCATGCACTAAGAAAGATGTTGAAGCGCTTCAAGAGCTTCAAGAGTCTTCAAGAGCGC 1320
 QY 1321 TGCAGATTTATGAAGCTGCGACACATG--GAAGACCTTCGTGACGCTGCAAGATGCT 1378
 DB 1321 TGCAGATTTATGAAGCTGCGACACATGGAAGAAAGCCCTGTCGAGCTGCAAGATGCT 1380
 QY 1379 GATGACACTGCGCACTCCTGAGGAGACCTTACCAAGGCGCTGACAGATTTCTAACAAT 1438
 DB 1379 GATGACACTGCGCACTCCTGAGGAGACCTTACCAAGGCGCTGACAGATTTCTAACAAT 1440
 QY 1439 CAAACTAGAAGGCAAGGCTCCATGACAAACTTTTGTGAATGTTGAGGCGCAAGGT 1498
 DB 1441 CAAACTAGAAGGCAAGGCTCCATGACAAACTTTTGTGAATGTTGAGGCGCAAGGT 1500
 QY 1499 CTGACTTAAGAGCTCCCTGAGGCTTCCCATCTTCAATGTTGAAGAAAGGAAATTAACCA 1558

Db	1501	CTGACTAAAGCTCCCTGGGCGCTTCCCATCTTCACTGTTGAAAAGGAAAAATAAACCA	1560
Qy	1559	AGAGTGAATGTCGAAAGAACTTAGAGTTAGTTAAACAACATCAAAAAATCAACAGACTGCAC	1618
Db	1561	AGAGTGAATGTCGAAAGAACTTAGAGTTAGTTAAACAACATCAAAAAATCAACAGACTGCAC	1620
Qy	1619	TGATATATTAGCAGCAAGACTTAGAGAGAGCTTCAGATTCCTCCATAGGTTCCATATA	1678
Db	1621	TGATATATTAGCAGCAAGACTTAGAGAGAGCTTCAGATTCCTCCATAGGTTCCATATA	1680
Qy	1679	GTCTCTCTCACTTCCACATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1738
Db	1681	GTCTCTCTCACTTCCACATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1740
Qy	1739	TATTTTTCTCCCTTTCTCTTTCACCTCCCTTAATTTCTTGGCTTCTTCAATCTAGTT	1798
Db	1741	TATTTTTCTCCCTTTCTCTTTCACCTCCCTTAATTTCTTGGCTTCTTCAATCTAGTT	1800
Qy	1799	CCCATCTCCCTTAATTTTCTTCCCGCTGCTGCTGCTCTTCTTCTTCTTACTCACTCC	1858
Db	1801	CCCATCTCCCTTAATTTTCTTCCCGCTGCTGCTGCTCTTCTTCTTCTTACTCACTCC	1860
Qy	1859	ATTCCCTCTTTCTCCATCCCTTCCCGCTTTTCTCAAAATTTGAAAATAGCTTAGTTTAAA	1918
Db	1861	ATTCCCTCTTTCTCCATCCCTTCCCGCTTTTCTCAAAATTTGAAAATAGCTTAGTTTAAA	1920
Qy	1919	AAAAAAATCCCTCCCTTCCCGCTTCTCTTCCCTTCTTCCCTTCTTCCCTTCTTCC	1978
Db	1921	AAAAAAATCCCTCCCTTCCCGCTTCTCTTCCCTTCTTCCCTTCTTCCCTTCTTCC	1980
Qy	1979	CTTTCTCTTCTCTTCTCTTGAACCTTCTTTCATCTTCTTCTTCTTCTTCTCTGCTG	2038
Db	1981	CTTTCTCTTCTCTTCTCTTGAACCTTCTTTCATCTTCTTCTTCTTCTTCTCTGCTG	2040
Qy	2039	AACTTTTAAAGAGTCTCTAACTGAAGAGATGGAAGCCAGCCCTGCCAAGAGATGA	2098
Db	2041	AACTTTTAAAGAGTCTCTAACTGAAGAGATGGAAGCCAGCCCTGCCAAGAGATGA	2100
Qy	2099	GATCCATATATGAGATGCGACAGTGAACCTATTTGTAACCATACCGTCCCAATGACTAAG	2158
Db	2101	GATCCATATATGAGATGCGACAGTGAACCTATTTGTAACCATACCGTCCCAATGACTAAG	2160
Qy	2159	AATCAAGAGAGAGAACCAACGTTCTTAAAGTACGTGCAACATATACAAATTTGACTGA	2218
Db	2161	AATCAAGAGAGAGAACCAACGTTCTTAAAGTACGTGCAACATATACAAATTTGACTGA	2220
Qy	2219	GTCAGATTTAGATTTCAATGGAGAGAGCTCTAATTAACAACCTTAAGCAACGTTGATC	2278
Db	2221	GTCAGATTTAGATTTCAATGGAGAGAGCTCTAATTAACAACCTTAAGCAACGTTGATC	2280
Qy	2279	GGCTCTCTCTAATCAATGCTTTTCCATCTAGTACAGTACAGCAATTTGATTCCTTAAT	2338
Db	2281	GGCTCTCTCTAATCAATGCTTTTCCATCTAGTACAGTACAGCAATTTGATTCCTTAAT	2340
Qy	2339	GTTTTTTCAAGTCTTCAAGATATTTGTTAGTTAGTACTATGTACTTTTTCAGGAAAT	2398
Db	2341	GTTTTTTCAAGTCTTCAAGATATTTGTTAGTTAGTACTATGTACTTTTTCAGGAAAT	2400
Qy	2399	AGTTTAAAGCTTAATTCATTCAGCAACTCTAAAGAAATAGAAATACTGCAATTTGCTG	2458
Db	2401	AGTTTAAAGCTTAATTCATTCAGCAACTCTAAAGAAATAGAAATACTGCAATTTGCTG	2460
Qy	2459	CTGGCTTTCGAACAATTTACGAACAATATAGAGAACCAATGATCTCGAAGAGAAATTTT	2518
Db	2461	CTGGCTTTCGAACAATTTACGAACAATATAGAGAACCAATGATCTCGAAGAGAAATTTT	2520
Qy	2519	AAAAATGTTTTGTTTTCTTCTCAAAATGAGATTTTTTTTGACAGCTTTTACCACTTTTC	2578
Db	2521	AAAAATGTTTTGTTTTCTTCTCAAAATGAGATTTTTTTTGACAGCTTTTACCACTTTTC	2580
Qy	2579	AGCCATTATTAATGGAATTTTAACTTACAGCAATAGTTGAAGGAAGGTGCATA	2638

CC having lipid metabolism regulatory and fatty acid synthesis functions.
CC The product of the invention has cytostatic and antiarteriosclerotic
CC activity. The protein and gene are useful as drugs and in developing
CC drugs for treatment of e.g. arteriosclerosis and diabetes. The protein is
CC obtained by cloning human brain-originated cDNA library. This sequence
CC encodes the human brain ERgamma protein which is described in the method
CC of the invention
XX
XX

Sequence 5216 BP; 1568 A; 1019 C; 974 G; 1655 T; 0 U; 0 Other;

Query Match 97.2%; Score 2901.8; DB 3; Length 5216;
Best Local Similarity 99.7%; Pred. No. 0;

Matches 2929; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

```
QY 17 GTGAATTCGGCTTGTGCTAGAGAAACATTTGTATTGCACTGTGCTGTGCAAG 76
D 46 GTGCTTATAGGCTGTGCTAGAGAAACATTTGTATTGCACTGTGCTGTGCAAG 105
QY 77 AAATTTGATTATAGCTGGGGTGCACAAATATGTTGCCGTGCAATGATTTGGT 136
D 106 AAATCTTGATTATAGCTGGGGTGCACAAATATGTTGCCGTGCAATGATTTGGT 165
QY 137 AGAATTTGCTTCTGCTGATCTTTTCCCTGCACTAGAGAGAGCTTCTGTGAGAT 196
D 166 AGAATTTGCTTCTGCTGATCTTTTCCCTGCACTAGAGAGAGCTTCTGTGAGAT 225
QY 197 GTCAACAAAGATGACATGATTTCCAGCTGTTCCTTCATCAAGAGAACTTC 256
D 226 GTCAACAAAGATGACATGATTTCCAGCTGTTCCTTCATCAAGAGAACTTC 285
QY 257 CAGCCAGCCCTCCCTGAGGACAGGTCACACCAAGCCCTGTGCTTTTCAAGGC 316
D 286 CAGCCAGCCCTCCCTGAGGACAGGTCACACCAAGCCCTGTGCTTTTCAAGGC 345
QY 317 CAGTGGAGCTACAGTTCAACATGAAATGSCATCAGACGAGCTTGACCTGCCACT 376
D 346 CAGTGGAGCTACAGTTCAACATGAAATGSCATCAGACGAGCTTGACCTGCCACT 405
QY 377 CTACCTTTCTGCTCTTACTCTGAGAGGTAGTGGCTGTGCAAGAACTGTATGATG 436
D 406 CTACCTTTCTGCTCTTACTCTGAGAGGTAGTGGCTGTGCAAGAACTGTATGATG 465
QY 437 CTCCAGACCATTTGTGAAGATCCCGACGACAGATGTAATGCTCACTGATGACC 496
D 466 CTCCAGACCATTTGTGAAGATCCCGACGACAGATGTAATGCTCACTGATGACC 525
QY 497 CAAGAGACTGTGTTAGTGTGTGACATCGCTTCTGGGTACCACTATGGGTAGCAT 556
D 526 CAAGAGACTGTGTTAGTGTGTGACATCGCTTCTGGGTACCACTATGGGTAGCAT 585
QY 557 ATGTGAAGCTGTCAAGGATTTCTTAAAGAGACATTCAGGCAATATAGATACAGCTG 616
D 586 ATGTGAAGCTGTCAAGGATTTCTTAAAGAGACATTCAGGCAATATAGATACAGCTG 645
QY 617 CCCGACACGAAATGTAATGTGAATCACAAGCGACAGATTAATCCCTGCGAGGTTGCCG 676
D 646 CCCGACACGAAATGTAATGTGAATCACAAGCGACAGATTAATCCCTGCGAGGTTGCCG 705
QY 677 CTTTATGAAGTGTAAAAAGTGGGCACTGTAAGAAAGAGGGGTGGCTCTTGCACAGATAG 736
D 706 CTTTATGAAGTGTAAAAAGTGGGCACTGTAAGAAAGAGGGGTGGCTCTTGCACAGATAG 765
QY 737 TGGAGGTGTGGAGAAATGCAAGCGAGATGATGCGGAGAAAGCCCATACCTGAACC 796
D 766 TGGAGGTGTGGAGAAATGCAAGCGAGATGATGCGGAGAAAGCCCATACCTGAACC 825
QY 797 TCAGCTGTTTCAGCGACGCAAAAAGCCATATTAACAATGTCTCAATTTGTTGGTGGC 856
D 826 TCAGCTGTTTCAGCGACGCAAAAAGCCATATTAACAATGTCTCAATTTGTTGGTGGC 885
QY 857 TGAACCGAGAAATCTATGCACTGACCTACTGCTCCCGACAGTGCATCAAAAGC 916
D 886 TGAACCGAGAAATCTATGCACTGACCTACTGCTCCCGACAGTGCATCAAAAGC 945
```

```
QY 917 CTTGACACACTGTGTGACTTGGCCGACCGAGAGTGGTGTATTCATTTGATGGCCGAA 976
D 946 CCTCAGTACACTGTGTGACTTGGCCGACCGAGAGTGGTGTATTCATTTGATGGCCGAA 1005
QY 977 GCATTTCCAGGCTTCTCCAGCTGTCTCCCTGGCGGACCAATGAGCTTCTGCAAGTGC 1036
D 1006 GCATTTCCAGGCTTCTCCAGCTGTCTCCCTGGCGGACCAATGAGCTTCTGCAAGTGC 1065
QY 1037 TTGATGAAATTTTGAATCCTGTGTGTGTATGACCGGTCTCTTCAATTTGAGATGAACT 1096
D 1066 TTGATGAAATTTTGAATCCTGTGTGTGTATGACCGGTCTCTTCAATTTGAGATGAACT 1125
QY 1097 TGTCTATGACAGCATTTATATGATGAGAGACAGTCCAAATTAGAGGCTCTTCTGA 1156
D 1126 TGTCTATGACAGCATTTATATGATGAGAGACAGTCCAAATTAGAGGCTCTTCTGA 1185
QY 1157 TCTTAATATGCTATCTCTGACGCTGTATGAATATCAAGCATGAAAGCTGGAAGAA 1216
D 1186 TCTTAATATGCTATCTCTGACGCTGTATGAATATCAAGCATGAAAGCTGGAAGAA 1245
QY 1217 AGAATTTGTCACTCTCAAGCTATAGCTCTTGTATTCAGACTCCATGCAATAGAGA 1276
D 1246 AGAATTTGTCACTCTCAAGCTATAGCTCTTGTATTCAGACTCCATGCAATAGAGA 1305
QY 1277 TGTGAAGCCGTTGAGAGCTTCAAGATGTCTTATCATGAAGCGCTGACAGATTATGAAGC 1336
D 1306 TGTGAAGCCGTTGAGAGCTTCAAGATGTCTTATCATGAAGCGCTGACAGATTATGAAGC 1365
QY 1337 TGGCAGACATATGGAAGACCTGTGTGAGCTGGAAGATGTGTATGACACTGCCACTCT 1396
D 1366 TGGCAGACATATGGAAGACCTGTGTGAGCTGGAAGATGTGTATGACACTGCCACTCT 1425
QY 1397 GAGCAGACCTCTTCCAAAGCCGTGACGACTTTTCAACAATCAAACTAGAAGCGAAAGT 1456
D 1426 GAGCAGACCTCTTCCAAAGCCGTGACGACTTTTCAACAATCAAACTAGAAGCGAAAGT 1485
QY 1457 CCGAATGCACAAACTTTTGTGAATGTGTGAAGCCCAAGCTGTGACTTAAAGCTCCCTG 1516
D 1486 CCGAATGCACAAACTTTTGTGAATGTGTGAAGCCCAAGCTGTGACTTAAAGCTCCCTG 1545
QY 1517 GGCCTTCCCATCTTCAATGTGTGAAGAAAGGAAATTAACCCAAAGATGTGCAAGAA 1576
D 1546 GGCCTTCCCATCTTCAATGTGTGAAGAAAGGAAATTAACCCAAAGATGTGCAAGAA 1605
QY 1577 CTTAGATTATGTTAACAACATCAAAATCAACAGACTGCACTGATATTTAGACAGCAG 1636
D 1606 CTTAGATTATGTTAACAACATCAAAATCAACAGACTGCACTGATATTTAGACAGCAG 1665
QY 1637 ACTATGAAGACTTTCAGATTCCTGCATAGAGTCTGTAGAG-TTCTTTCTACTTCTC 1695
D 1666 ACTATGAAGACTTTCAGATTCCTGCATAGAGTCTGTAGAGTCTTCTACTTCTC 1725
QY 1696 CATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1755
D 1726 CATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1785
QY 1756 TTCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1815
D 1786 TTCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1845
QY 1816 TCTTCCGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1875
D 1846 TCTTCCGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1905
QY 1876 TCTTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1935
D 1906 TCTTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1963
QY 1936 CCCCTTTCCTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1995
D 1964 CCCCTTTCCTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2023
```

QY 1996 CTTGACCTCTTCCATCTTTCTTTCTCTGCTGCTGAACCTTTTAAAGAGTC 2055
 DB 2024 CTTGACCTCTTCCATCTTTCTTTCTCTGCTGCTGAACCTTTTAAAGAGTC 2083
 QY 2056 TCTAACTGAAGAGATGAGAGCCAGCCCTGCCAAAGATGAGATCCATTAATGATG 2115
 DB 2084 TCTAACTGAAGAGATGAGAGCCAGCCCTGCCAAAGATGAGATCCATTAATGATG 2143
 QY 2116 CCAGTGAACCTTATGTAACCAATCCGTCCTCCCAATCACTTAAGATCAAGAGAGAAC 2175
 DB 2144 CCAGTGAACCTTATGTAACCAATCCGTCCTCCCAATCACTTAAGATCAAGAGAGAAC 2203
 QY 2176 CAACGTTCTTAAAGATGAGATGAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2235
 DB 2204 CAACGTTCTTAAAGATGAGATGAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2263
 QY 2236 AAGGAGAGAGCTCTTAATTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2295
 DB 2264 AAGGAGAGAGCTCTTAATTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2323
 QY 2296 GCTTTCCATTAAGATCACTTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2355
 DB 2324 GCTTTCCATTAAGATCACTTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2383
 QY 2356 AAGTATTTGTTAGTAACTTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2415
 DB 2384 AAGTATTTGTTAGTAACTTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2443
 QY 2416 TTGATCAATTAAGATCACTTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2475
 DB 2444 TTGATCAATTAAGATCACTTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2503
 QY 2476 CGAACAATTAAGATCACTTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2535
 DB 2504 CGAACAATTAAGATCACTTAAGATCACTTAAGATCACTTAAGATCACTTAAGAT 2563
 QY 2536 TCTTAACTGAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2595
 DB 2564 TCTTAACTGAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2623
 QY 2596 GGAATTTAACTTAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2655
 DB 2624 GGAATTTAACTTAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2683
 QY 2656 TATGTTGTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2715
 DB 2684 TATGTTGTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2743
 QY 2716 AATGTTGTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2775
 DB 2744 AATGTTGTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2803
 QY 2776 ATTAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2835
 DB 2804 ATTAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2863
 QY 2836 GAGAGTAACTTGAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2895
 DB 2864 GAGAGTAACTTGAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2923
 QY 2896 GTTTGCAAAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2954
 DB 2924 GTTTGCAAAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 2982

RESULT 6
 ADD69756
 ID ADD69756 standard; DNA; 5216 BP.
 XX
 AC
 XX ADD69756;
 DT 15-JAN-2004 (first entry)
 XX

DE Human ERR gamma 3-related DNA - SEQ ID 5.
 XX
 XX nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;
 KW oestrogen receptor; ER; thyroid hormone; TR; human; ds; gene.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 155..1531
 FT /tag a
 FT /product= "Human ERR gamma 3-related protein - SEQ ID 6"
 PN WO2003080831-A1.
 PD 02-OCT-2003.
 XX
 XX 25-MAR-2003; 2003WO-UP003611.
 XX
 XX 25-MAR-2002; 2002JP-00084560.
 PR
 XX
 PA (FUJII) FUJISAWA PHARM CO LTD.
 XX
 XX Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;
 DR
 DR MPI; 2003-779262/73.
 XX
 XX P-PSDB; ADD69757.
 PT Polynucleotides encoding nuclear receptors, and the encoded proteins,
 PT useful as diagnostic agents, and for identification of agents that affect
 PT receptor activity.
 PS
 PS Claim 1; SEQ ID NO 5; 148bp; Japanese.
 XX
 XX The invention relates to novel nuclear receptor ERR (oestrogen receptor-
 CC related receptor) gamma 3 polynucleotides. The polynucleotides of the
 CC invention may be useful for diagnosis of disorders caused by abnormal
 CC nuclear receptor activity, particularly those related to abnormal
 CC oestrogen receptor (ER), ER or thyroid hormone receptor (TR) activity.
 CC Furthermore, the polynucleotides and proteins may be useful for
 CC evaluating agents that affect the activity of nuclear receptors. The
 CC current sequence is that of the human ERR gamma 3-related DNA of the
 CC invention.
 XX
 XX Sequence 5216 BP; 1568 A; 1019 C; 974 G; 1655 T; 0 U; 0 Other;
 SQ
 Query Match 97.2%; Score 2901.8; DB 9; Length 5216;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2929; Conservative 0; Mismatches 7; Indels 3; Gaps 2;
 QY 17 GTGGAATTCGGCTGTCTAGAGAGAACTTTGTGTTAATGCACTGTGCTGTCAAGG 76
 DB 46 GTGCTATAGGGCTGTGCTAGAGAGAACTTTGTGTTAATGCACTGTGCTGTCAAGG 105
 QY 77 AATCTTGAATTAATGCTGGGGTGCACAAATATATGTTGCGCGGTCCGACATGATTCGT 136
 DB 106 AATCTTGAATTAATGCTGGGGTGCACAAATATATGTTGCGCGGTCCGACATGATTCGT 165
 QY 137 AGAATTTGGCTTCCGAATCTTTTCCCTGCACTACAGAGAGAGCTTCTGCAAGAT 196
 DB 166 AGAATTTGGCTTCCGAATCTTTTCCCTGCACTACAGAGAGAGCTTCTGCAAGAT 225
 QY 197 GTCAAACAAAAGATGACATTTGATTCAGCTGTGCTCTTCAATCAAGAGAACTTC 256
 DB 226 GTCAAACAAAAGATGACATTTGATTCAGCTGTGCTCTTCAATCAAGAGAACTTC 285
 QY 257 CAGCCAGCTCCCTGACGAGAGAGCTTCAACCAACCAAGCCCTGTGCTCTTCAAGAG 316
 DB 286 CAGCCAGCTCCCTGACGAGAGAGCTTCAACCAACCAAGCCCTGTGCTCTTCAAGAG 345
 QY 317 CAGTGGAGCTACAGTTCAACCATGAAATGAGCAATCAAGAGAGCTTCAAGAGAGCTT 376
 DB 346 CAGTGGAGCTACAGTTCAACCATGAAATGAGCAATCAAGAGAGCTTCAAGAGAGCTT 405

QY 377 CTACCTTCTGCTCTCTATCTGAGAGGTAGTGGGCTGTCAAGAACTGTATGATGACTG 436
Db 406 CTACCTTCTGCTCTCTATCTGAGAGGTAGTGGGCTGTCAAGAACTGTATGATGACTG 465
QY 437 CTCAGCAGCATTTGTGAGAGTCCCGAGACCAAGTGTGATACATGCTCACTCGATGCC 496
Db 466 CTCAGCAGCATTTGTGAGAGTCCCGAGACCAAGTGTGATACATGCTCACTCGATGCC 525
QY 497 CAAGAGACTGTGTGTAGTGTGTGTGACATCGCTTCTGGGTACCACTATGGGTAGCATC 556
Db 526 CAAGAGACTGTGTGTAGTGTGTGTGACATCGCTTCTGGGTACCACTATGGGTAGCATC 585
QY 557 ATGTGAAGCCTGCAAGGCATTTCTTCAAGAGACATTTCAAGGCATTTAAGATACAGCTG 616
Db 586 ATGTGAAGCCTGCAAGGCATTTCTTCAAGAGACATTTCAAGGCATTTAAGATACAGCTG 645
QY 617 CCCGCGCAGCAATGATGTGAAATCACAAGCGCAGAGCTPAATCCGCGCAGGCTTSCCG 676
Db 646 CCCGCGCAGCAATGATGTGAAATCACAAGCGCAGAGCTPAATCCGCGCAGGCTTSCCG 705
QY 677 CTTTCATGAAGTGTTTAAAGTGGGCATGCTGAAGAAAGGGGTGCGCTTGAACAGATACG 736
Db 706 CTTTCATGAAGTGTTTAAAGTGGGCATGCTGAAGAAAGGGGTGCGCTTGAACAGATACG 765
QY 737 TGGAGGTGCGGAGAGTACAGAGCCGAGATAGATGCGGAGACACGCCATACCTGAACC 796
Db 766 TGGAGGTGCGGAGAGTACAGAGCCGAGATAGATGCGGAGACACGCCATACCTGAACC 825
QY 797 TCAGCTGGTTCAGCGCAGCAGCAAAAAGCCATTAACAAGTTGTCTCACATTTGTTGGTGGC 856
Db 826 TCAGCTGGTTCAGCGCAGCAGCAAAAAGCCATTAACAAGTTGTCTCACATTTGTTGGTGGC 885
QY 857 TGAACCGAGAGAGATCTATGCGCATGCGTGAACCTTACTGTCCCGCAGAGTGAATCAAGC 916
Db 886 TGAACCGAGAGAGATCTATGCGCATGCGTGAACCTTACTGTCCCGCAGAGTGAATCAAGC 945
QY 917 CTTCACTACACTGTGTGACTTTGGCCGACCGAGAGTGTGTGTATCATTTGATGGGAGAA 976
Db 946 CTTCACTACACTGTGTGACTTTGGCCGACCGAGAGTGTGTGTATCATTTGATGGGAGAA 1005
QY 977 GCATATTCAGAGGCTTTCACAGCGGTGCTGAGCGGAGACCAAGTGAAGCTTTCAGAGATGC 1036
Db 1006 GCATATTCAGAGGCTTTCACAGCGGTGCTGAGCGGAGACCAAGTGAAGCTTTCAGAGATGC 1065
QY 1037 TTGGATGAAATTTTGAATCCTTGGTGTGTATACCGGTCTCTTTCATTTGAGATGAAT 1096
Db 1066 TTGGATGAAATTTTGAATCCTTGGTGTGTATACCGGTCTCTTTCATTTGAGATGAAT 1125
QY 1097 TGTCTATGACAGCATTTATATATATGACGAAAGACCACTCCAAATTTAGAGGCTTCTTGA 1156
Db 1126 TGTCTATGACAGCATTTATATATATGACGAAAGACCACTCCAAATTTAGAGGCTTCTTGA 1185
QY 1157 TCTPAATAATGCTATCCCTGAGCTGTGTAAAGAAATACAAAGCATGGAAGCTGGAAGAA 1216
Db 1186 TCTPAATAATGCTATCTGAGCTGTGTAAAGAAATACAAAGCATGGAAGCTGGAAGAA 1245
QY 1217 AGAATTTGTCAACCTCAAGGCTATAGCTCTTGTCTAATTCAGACTCCATGCAATPAGAA 1276
Db 1246 AGAATTTGTCAACCTCAAGGCTATAGCTCTTGTCTAATTCAGACTCCATGCAATPAGAA 1305
QY 1277 TGTGTGAAGCCGTTTACGAAGCTTCAGGATGTCTTACATGAAGCGCTGCAAGATTTATGAAGC 1336
Db 1306 TGTGTGAAGCCGTTTACGAAGCTTCAGGATGTCTTACATGAAGCGCTGCAAGATTTATGAAGC 1365
QY 1337 TGGCCACACATGGAAGACCTCGTCGAGCTGCGAAGATGTGATGACACTGCCACTCCT 1396
Db 1366 TGGCCACACATGGAAGACCTCGTCGAGCTGCGAAGATGTGATGACACTGCCACTCCT 1425
QY 1397 GAGGCAAGCTCTTACCAAGGCGGTGACGATTTCTTCAACATCAAACTPAGAGGCAAGT 1456
Db 1426 GAGGCAAGCTCTTACCAAGGCGGTGACGATTTCTTCAACATCAAACTPAGAGGCAAGT 1485
QY 1457 CCCAATGACAAACTTTTGTGAAATGTGTGAGGCCAAGTCTGACTAAAGACTCCCTG 1516

Db 1486 CCCAATGACAAACTTTTGTGAAATGTGTGAGGCCAAGTCTGACTAAAGCTCCCTG 1545
QY 1517 GGCCTTCCCATCTCTTCATGTGTGAAAAAGGAAAAATTAACCAAGAGTATGAGAA 1576
Db 1546 GGCCTTCCCATCTCTTCATGTGTGAAAAAGGAAAAATTAACCAAGAGTATGAGAA 1605
QY 1577 CTTAGAGTTTATTAACAACATCAAAATATCAACAGACTGCACTGATATATTTAGACGAA 1636
Db 1606 CTTAGAGTTTATTAACAACATCAAAATATCAACAGACTGCACTGATATATTTAGACGAA 1665
QY 1637 ACTATGAAGAGCTTTCAGATTCCTCGATAGGTTCCTGAGAG--TCTCTTCACTTCTC 1695
Db 1666 ACTATGAAGAGCTTTCAGATTCCTCGATAGGTTCCTGAGAGTTCTCTTCACTTCTC 1725
QY 1696 CATCATCTTCTTCTCTCTCTTCCACATTTCTCTTCTCTTCTTCTTCTTCTTCTTCT 1755
Db 1726 CATCATCTTCTTCTCTCTTCTTCCACATTTCTCTTCTCTTCTTCTTCTTCTTCT 1785
QY 1756 TCTTTTCACTCCCTTATTTCTTGTGCTTCTTCAATTCCTAGTCCATTTCTCTTATTT 1815
Db 1786 TCTTTTCACTCCCTTATTTCTTGTGCTTCTTCAATTCCTAGTCCATTTCTCTTATTT 1845
QY 1816 TCTTCCGCTGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1875
Db 1846 TCTTCCGCTGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1905
QY 1876 TCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1935
Db 1906 TCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1963
QY 1936 CCCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1995
Db 1964 CCCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2023
QY 1996 CTTGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2055
Db 2024 CTTGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2083
QY 2056 TCTPACTGAAGAGATGAGAGCCAGCCGCAAGAGATGAGATCCATTAATGATG 2115
Db 2084 TCTPACTGAAGAGATGAGAGCCAGCCGCAAGAGATGAGATCCATTAATGATG 2143
QY 2116 CCAAGTGAATTTATTTGAAACCATACCTGTCCCAATGACTAAGAGATCAAGAGAGAAC 2175
Db 2144 CCAAGTGAATTTATTTGAAACCATACCTGTCCCAATGACTAAGAGATCAAGAGAGAAC 2203
QY 2176 CAAGCTTCTTAAAGTACAGTGAACATTAACAATGACAGAGTGAATTTGATTTG 2235
Db 2204 CAAGCTTCTTAAAGTACAGTGAACATTAACAATGACAGAGTGAATTTGATTTG 2263
QY 2236 ATGGAGACAGCTCTTATTTAGACAACCTTAAGCAAGTTGCACTGGCTCTTATCATTT 2295
Db 2264 ATGGAGACAGCTCTTATTTAGACAACCTTAAGCAAGTTGCACTGGCTCTTATCATTT 2323
QY 2296 GCTTTTCATCTAGATGATTAACGCAATTTGATTTCTTATTTTCTTCAAGCTTCTC 2355
Db 2324 GCTTTTCATCTAGATGATTAACGCAATTTGATTTCTTATTTTCTTCAAGCTTCTC 2383
QY 2356 AGGTATTTGTATGTTAGCTTACATATGTAACCTTTTCAAGGAAATGTTTAACTTATTTCA 2415
Db 2384 AGGTATTTGTATGTTAGCTTACATATGTAACCTTTTCAAGGAAATGTTTAACTTATTTCA 2443
QY 2416 TTTCAATGCAATTAAGAGAAATTAAGAAATCTGCAATTTTGTGCTTGAACAATTA 2475
Db 2444 TTTCAATGCAATTAAGAGAAATTAAGAAATCTGCAATTTTGTGCTTGAACAATTA 2503
QY 2476 CGAACAATTAATGAAGCAAAATGATCTGGAAGAAATTTTAAATATGTTTGTCTTCT 2535
Db 2504 CGAACAATTAATGAAGCAAAATGATCTGGAAGAAATTTTAAATATGTTTGTCTTCT 2563
QY 2536 TCTTCAAAATGGAATTTTGTGACAGCTTTTCAACTTTTCAAGCTTATTAATATG 2595

Db 1101 CTTTCGTTGAGATGTAACCTTCTATGACAGCATTTATATGACGAGCCAGTGC 1160
 QY 1137 AAATGACAGGCGCTTCTGTATCTAAATATATGCTATCTGACGTGTAAAGAAATCAAG 1196
 Db 1161 AAATGACAGGCGCTTCTGTATCTAAATATATGCTATCTGACGTGTAAAGAAATCAAG 1220
 QY 1197 AGATGAGGTGAAAAAGAGAAATTTGTGACCCCTCAAGACTATAGTCTTGTCTAATTC 1256
 Db 1221 AGCATGAGGTGAAAAAGAGAAATTTGTGACCCCTCAAGACTATAGTCTTGTCTAATTC 1280
 QY 1257 GACTCCATGACATATGAGATGTTGAAGCGCTCAGAGCTTCAAGATGCTTACATGAA 1316
 Db 1281 GACTCCATGACATATGAGATGTTGAAGCGCTTCAAGACTTCAAGATGCTTACATGAA 1340
 QY 1317 GCGCTGACAGATTTATGAGCTGCGCAGACATGAGAAACCCCTGTGAGCTGCAAGATG 1376
 Db 1341 GCGCTGACAGATTTATGAGCTGCGCAGACATGAGAAACCCCTGTGAGCTGCAAGATG 1400
 QY 1377 CTGATGACATGCTGCTCTGAGAGCAGACTCTACCAAGGCGTGCAGCATTTCTACAC 1436
 Db 1401 CTGATGACATGCTGCTCTGAGAGCAGACTCTACCAAGGCGTGCAGCATTTCTACAC 1460
 QY 1437 ATCAAACTGAGAGCAAAAGTCCCAATGACACAACTTTTGGAAAATGTTGAGAGCCAG 1496
 Db 1461 ATCAAACTGAGAGCAAAAGTCCCAATGACACAACTTTTGGAAAATGTTGAGAGCCAG 1520
 QY 1497 GTCTGACTAAAGCTCCCTGCGGCTTCCCATCTTCATGTTGAAAAAGGAAATTAACC 1556
 Db 1521 GTCTG-CTAAAGCTCCCTGCGGCTT-CCATCTTCATGTTGAAAAAGGAAATTAACC 1578
 QY 1557 CAGAGGTGATGAGAGAACTTAGATTTAGTTAACAATCAAAATCAACAGACTGC 1616
 Db 1579 CAGAGGTGATGAGAGAACTTAGATTTAGTTAACAATCAAAATCAACAGACTGC 1638
 QY 1617 ACTGATATTTAGAGCAAGACTATGAGAGCGTTTCAAGATCCCTCATAGGTTCCGAT 1676
 Db 1639 ACTGATATTTAGAGCAAGACTATGAGAGCGTTTCAAGATCCCTCATAGGTTCCGAT 1698
 QY 1677 GAG-TTCTTCTAATCTTCTCATCATCTTCTTCTCTTCTTCTTCCACATTTCTCTTCT 1735
 Db 1699 GAGTTCTTCTTCTTCTTCTCATCATCTTCTTCTTCTTCTTCCACATTTCTCTTCT 1758
 QY 1736 CTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1795
 Db 1759 CTTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1818
 QY 1796 GTTCCCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1855
 Db 1819 GTTCCCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1878
 QY 1856 CTGATTCCTCTCTTCTCTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1915
 Db 1879 CTGATTCCTCTCTTCTCTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1937
 QY 1916 AAAAAAAAAATCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1975
 Db 1938 -AAAAAAAAATCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1996
 QY 1976 TCCCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2035
 Db 1997 TCCCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2066
 QY 2036 CTGAACCTTTTAAAGAGTCTCTAATGAGAGAGATGAGAGCCAGCCCTGCAAGAGAT 2095
 Db 2057 CTGAACCTTTTAAAGAGTCTCTAATGAGAGAGATGAGAGCCAGCCCTGCAAGAGAT 2116
 QY 2096 GGAATTCATATATGATGAGTCACTGAACTTATGTAACCATACCGTCCCATATGACTA 2155
 Db 2117 GGAATTCATATATGATGAGTCACTGAACTTATGTAACCATACCGTCCCATATGACTA 2176
 QY 2156 AGAATCAAAAGAGAGAGCAACGTTCTTAAAGTACAGTCAACATATCAATATGAC 2215
 Db 2177 AGAATCAAAAGAGAGAGCAACGTTCTTAAAGTACAGTCAACATATCAATATGAC 2235

QY 2216 TGAATGACATATTTAGATTTTCATGAGAGAGCGCTCTTATTTAGACACTTAAGCAAGTTGC 2275
 Db 2236 TGAATGACATATTTAGATTTTCATGAGAGAGCGCTCTTATTTAGACACTTAAGCAAGTTGC 2295
 QY 2276 ATGGGCTGCTCTTATCATTTGCTTTTCCATCTAGATCAGTTACAGCCATTTATCTCTTA 2335
 Db 2296 ATGGGCTGCTCTTATCATTTGCTTTTCCATCTAGACAGTTACAGCCATTTATCTCTTA 2355
 QY 2336 ATGTTTTTTCAGTCTCCAGATATTTGTATTTAGTACTATGTAACTTTTTCAGGG 2395
 Db 2356 ATGTTTTTTCAGTCTCCAGATATTTGTATTTAGTACTATGTAACTTTTTCAGGG 2415
 QY 2396 AATGTTTTAGCTTTATTTCAATTCATGCAATCTTAAAGAGAAATATCTGCAATTT 2455
 Db 2416 AATGTTTTAGCTTTATTTCAATTCATGCAATCTTAAAGAGAAATATATGCAATTT 2475
 QY 2456 GTGCTGCTTTGAGCAATTAACGACATATATGAAAGCAAAATGATCTTGAAGAGAT 2515
 Db 2476 GTGCTGCTTTGAGCAATTAACGACATATATGAAAGCAAAATGATCTTGAAGAGAT 2535
 QY 2516 TTTAAAAATGTTTTGTTCTCTTCAAAATGAGATTTTTTTTTCACAGCTTTTCACTT 2575
 Db 2536 TTTAAAAATGTTTTGTTCTCTTCAAAATGAGATTTTTTTTTCACAGCTTTTCACTT 2595
 QY 2576 TTTAGGCTTTTATTTATATGAGAAATTTTAACTTCTAAGAAATGTTGAGAGAGTGC 2635
 Db 2596 TTTAGGCTTTTATTTATATGAGAAATTTTAACTTCTAAGAAATGTTGAGAGAGTGC 2655
 QY 2636 ATATTTACAGGATGACATTTTANG--TTGTGTGCGCAGTCTGTGTCGCAACATCAATTTG 2692
 Db 2656 ATATTTACAGGATGACATTTTANGTTGTGTGCGCAGTCTGTGTCGCAACATCAATTTG 2715
 QY 2693 TTTAATGAGCTTCAGTTTAACTTAAATGTTCACTGACACAAAGATGAGATTAACCTTAC 2752
 Db 2716 TTTAATGAGCTTCAGTTTAACTTAAATGTTCACTGACACAAAGATGAGATTAACCTTAC 2775
 QY 2753 AGTGACTCTGAGTACATATATATTAAGCACTGACATGAGATATATGATCCGTGAAATG 2812
 Db 2776 AGTGACTCTGAGTACATATATATTAAGCACTGACATGAGATATATGATCCGTGAAATG 2835
 QY 2813 TCAGAGGTGACCTCTCTACTTGGAGGTACAAATGTCATATGATTTCTGATGCAATG 2872
 Db 2836 TCAGAGGTGACCTCTCTACTTGGAGGTACAAATGTCATATGATTTCTGATGCAATG 2895
 QY 2873 TGGTTAGAAATGTAATAC-TGCCGTGTTGCAAAATGACAGACC-TTGCTTCAAGAGAGC 2930
 Db 2896 TGGTTAGAAATGTAATACATGCTGCTTGGCAAGTCAAGACATTTGCTCAGAGAGAGC 2955
 QY 2931 TGTGAGCAGATTTCAATTTAAGAG 2954
 Db 2956 TGTGAGCAGATTTCAATTTAAGAG 2979

RESULT 8
 ADD69752
 ID ADD69752 standard; cDNA; 3362 BP.
 XX
 ADD69752;
 AC
 XX
 DT 15-JAN-2004 (first entry)
 XX
 Human ERR gamma 3-related cDNA - SEQ ID 1.
 XX
 nucleic acid receptor; ERR gamma 3; oestrogen receptor-related receptor;
 KW oestrogen receptor; ER; thyroid hormone; TR; human; ss; gene.
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 FH 684..1874
 FT
 CDS
 FT
 FT /product= "Human ERR gamma 3-related protein - SEQ ID 2"

Best Local Similarity 99.7%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 17 GTGGAAATTCGGCTTTCACATAGAGAAACATTTGTGTAAATTGCACTGTCTCTGTCAAG 76
Db 49 GTGGCTAATAGGCTGTGTACATAGAGAAACATTTGTGTAAATTGCACTGTCTCTGTCAAG 108
QY 77 AAACCTTGAATTAATAGCTGGGGTGCACAAATATATGTTGCCGTGCACATGATTCGGT 136
Db 109 AAACCTTGAATTAATAGCTGGGGTGCACAAATATATGTTGCCGTGCACATGATTCGGT 168
QY 137 AGAATTTGCTCTTCTGTAAATCTTTTTCCTGTGCACTACAGAGAAAGCTTCTGTGAGAA 196
Db 169 AGAATTTGCTCTTCTGTAAATCTTTTTCCTGTGCACTACAGAGAAAGCTTCTGTGAGAA 228
QY 197 GTCAAAACAAAGATGACACATTTGATTCAGCTGTGTGCTTCTTCACTACAGAGAACTTTC 256
Db 229 GTCAAAACAAAGATGACACATTTGATTCAGCTGTGTGCTTCTTCACTACAGAGAACTTTC 288
QY 257 CAGCCAGCCTCTCTGAGCGACAGGTCACACACACAGCCTGTGGTCTTTCAGAGCC 316
Db 289 CAGCCAGCCTCTCTGAGCGACAGGTCACACACACAGCCTGTGGTCTTTCAGAGCC 348
QY 317 CAGTGGAGCTACAGATTCAACCATGAAATGCGCATGAAACGCACTTGTACTGTGCCACTT 376
Db 349 CAGTGGAGCTACAGATTCAACCATGAAATGCGCATGAAACGCACTTGTACTGTGCCACTT 408
QY 377 CTACCTTCTGTCTCTATCTCTGGAGGTATGAGGCTGTGAGAGAACTGTATGATGACTG 436
Db 409 CTACCTTCTGTCTCTATCTCTGGAGGTATGAGGCTGTGAGAGAACTGTATGATGACTG 468
QY 437 CTCACGACCATTTGTGAAGATCCCGACAGACCAAGTGTGAATACATGCTCAATCTCGATGCC 496
Db 469 CTCACGACCATTTGTGAAGATCCCGACAGACCAAGTGTGAATACATGCTCAATCTCGATGCC 528
QY 497 CAAGGACTGTGTATGATGTGTGTGTGATCGCTTCTGAGTACCACTATGAGGATGACATC 556
Db 529 CAAGGACTGTGTATGATGTGTGTGTGATCGCTTCTGAGTACCACTATGAGGATGACATC 588
QY 557 ATGTGAACCTGTGACAGGCAATTTCAAGAGGCAATTTCAAGGCAATATAGAAATACAGCTG 616
Db 589 ATGTGAACCTGTGACAGGCAATTTCAAGAGGCAATTTCAAGGCAATATAGAAATACAGCTG 648
QY 617 CCTGTGCACGATGATGTGAATCAAAAGCGCAGAGCTGAATCTGTCCAGGCTTGTGCCG 676
Db 649 CCTGTGCACGATGATGTGAATCAAAAGCGCAGAGCTGAATCTGTCCAGGCTTGTGCCG 708
QY 677 CTTCAATGAAGTGTAAAAGTGGGATGTCTGAAGAAAGGGGTGCGTCTTTGACAGATACG 736
Db 709 CTTCAATGAAGTGTAAAAGTGGGATGTCTGAAGAAAGGGGTGCGTCTTTGACAGATACG 768
QY 737 TGAAGGTGCGAGAGATCAAGACGCGAGATAGATCGGAGAAACAGCCCATCTGAAACC 796
Db 769 TGAAGGTGCGAGAGATCAAGACGCGAGATAGATCGGAGAAACAGCCCATCTGAAACC 828
QY 797 TGAAGTGTGACGACGCAAAAGCCATATTAACAAGATGTCTCAATTTGTGTGTGCG 856
Db 829 TGAAGTGTGACGACGCAAAAGCCATATTAACAAGATGTCTCAATTTGTGTGTGCG 888
QY 857 TGAACCGGAGAAAGATCTATGCTATGCTGACCTGTCTCTCCCGCAGATGACATTAAGC 916
Db 889 TGAACCGGAGAAAGATCTATGCTATGCTGACCTGTCTCTCCCGCAGATGACATTAAGC 948
QY 917 CCTCACTACACTGTGTGACTTGTGCGACGACGAGAGTGTGTGTATCATTTGATGGCGAA 976
Db 949 CCTCACTACACTGTGTGACTTGTGCGACGACGAGAGTGTGTGTATCATTTGATGGCGAA 1008
QY 977 GCATATTCAGAGCTTCTCACGCTGTCTCTGTGCGACCAAGATGAGCTTCTGACAGAGTGC 1036
Db 1009 GCATATTCAGAGCTTCTCACGCTGTCTCTGTGCGACCAAGATGAGCTTCTGACAGAGTGC 1068
QY 1037 TTGATATGAATTTTGAATCTTGTGTGTGTATCCGCTCTCTTCAATTTGAGATGAATC 1096

```

```

Db 1069 TTGATATGAATTTTGAATCTTGTGTGTGTATCCGGTCTCTTCAATTTGAGATGAATC 1128
QY 1097 TGTCTATGACAGCATTTATATATATGACAGAACCAATGCTCAAAATTTGACAGGCTTCTTGA 1156
Db 1129 TGTCTATGACAGCATTTATATATATGACAGAACCAATGCTCAAAATTTGACAGGCTTCTTGA 1188
QY 1157 TCTAATATATGCTATCTCCGACAGCTGTAAAGAAATTAACAAGACATGAACTGGAAAAA 1216
Db 1189 TCTAATATATGCTATCTCCGACAGCTGTAAAGAAATTAACAAGACATGAACTGGAAAAA 1248
QY 1217 AGAATTTGTACCCCTCAAGCTATAGCTCTTGTCAATTCAGACTCCATGCACATAGAGAA 1276
Db 1249 AGAATTTGTACCCCTCAAGCTATAGCTCTTGTCAATTCAGACTCCATGCACATAGAGAA 1308
QY 1277 TGTGAAACCGCTTCAAGAGCTTCAAGATGTCTTAATCATGAAGCGCTGCAGATTTAGAAC 1336
Db 1309 TGTGAAACCGCTTCAAGAGCTTCAAGATGTCTTAATCATGAAGCGCTGCAGATTTAGAAC 1368
QY 1337 TGGCAGACATGGAAGACCCCTGTGAGCTGGCAAGATGTGATGACACTGCCACTCT 1396
Db 1369 TGGCAGACATGGAAGACCCCTGTGAGCTGGCAAGATGTGATGACACTGCCACTCT 1428
QY 1397 GAGCGACACTTCTACCAAGCCGTGACAGCATTTCTTAACAATCAAACTAGAAAGCAAA 1456
Db 1429 GAGCGACACTTCTACCAAGCCGTGACAGCATTTCTTAACAATCAAACTAGAAAGCAAA 1488
QY 1457 CCGAATGACAAACCTTTTGTGAAGTGTGAAGCCAAAGTCTCACTTAAGCTCCCTG 1516
Db 1489 CCGAATGACAAACCTTTTGTGAAGTGTGAAGCCAAAGTCTCACTTAAGCTCCCTG 1548
QY 1517 GGCCTTCCCATCTTCAATGTGAAGAAAGGAAATTAACCCCAAGATGATGCAAGAAA 1576
Db 1549 GGCCTTCCCATCTTCAATGTGAAGAAAGGAAATTAACCCCAAGATGATGCAAGAAA 1608
QY 1577 CTTAGAG 1583
Db 1609 CTTAGAG 1615

```

RESULT 10
AAZ95741
ID AAZ95741 standard; cDNA; 1431 BP.
XX
AC AAZ95741;
XX
DT 14-JUN-2000 (first entry)
XX
DE Human oestrogen related receptor 3 encoding cDNA SEQ ID NO:2.
XX
KW Human; ligand-combining; oestrogen related receptor; ERR3; diagnosis;
XX inflammation; cancer; osteoporosis; diabetes; renal disease; ss.
OS Homo sapiens.
XX
PN JP2000041681-A.
XX
PD 15-FEB-2000.
XX
PE 31-JUL-1998; 98JP-00217933.
XX
PR 31-JUL-1998; 98JP-00217933.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
WP1. 2000-295782/26.
DR P-PSDB; AAY82263.
XX
PT A new protein comprising a 435 amino acid sequence.
XX
PS Claim 6; Page 19-20; 22pp; Japanese.
XX
CC The present sequence encodes a human oestrogen related receptor 3 (ERR3)
protein which has ligand-combining activity. The polynucleotide sequence

CC encoding ERR3, and fragments of it, can be used for the diagnosis and the
 CC treatment of inflammations, cancers, osteoporosis, diabetes or renal
 CC diseases
 XX
 SQ Sequence 1431 BP; 401 A; 351 C; 348 G; 331 T; 0 U; 0 Other;
 Query Match 47.5%; Score 1418.2; DB 3; Length 1431;
 Best Local Similarity 99.8%; Pred. No. 1.1e-312;
 Matches 1420; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 176 GGAAGAGCTTCTCTGAGAGATGTCACAAAGATGCAACATGATGTCAGAGCTGTCGTC 235
 DB 9 GGAAGAGCTTCTCTGAGAGATGTCACAAAGATGCAACATGATGTCAGAGCTGTCGTC 68
 QY 236 CTTCATCAAGAGAGAGAGCTTCCAGCCAGCCCTCCCTGAGAGAGAGCTCAACCCAGCAG 295
 DB 69 CTTCATCAAGAGAGAGAGCTTCCAGCCAGCCCTCCCTGAGAGAGAGCTCAACCCAGCAG 128
 QY 296 CCTGTGCTCTCTTCAAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 355
 DB 129 CCTGTGCTCTCTTCAAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 188
 QY 356 CGGACTTGAGCTGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 415
 DB 189 CGGACTTGAGCTGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 248
 QY 416 CAGAGAACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 475
 DB 249 CAGAGAACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 308
 QY 476 ATACATGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 535
 DB 309 ATACATGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 368
 QY 536 GTACCCCTATGAG 595
 DB 369 GTACCCCTATGAG 428
 QY 596 AGGCAATATAGAAATACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
 DB 429 AGGCAATATAGAAATACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488
 QY 656 TAAATCTGCGAGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
 DB 489 TAAATCTGCGAGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 548
 QY 716 GGTGCGCTTGAAG 775
 DB 549 GGTGCGCTTGAAG 608
 QY 776 GAAAGAGCCATACCTGAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
 DB 609 GAAAGAGCCATACCTGAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 668
 QY 836 TGTCTCAATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
 DB 669 TGTCTCAATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728
 QY 896 CCGGAG 955
 DB 729 CCGGAG 788
 QY 956 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1015
 DB 789 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 848
 QY 1016 GATGAGCTTCTGAG 1075
 DB 849 GATGAGCTTCTGAG 908
 QY 1076 TCTTTCAATTTGAG 1135
 DB 909 TCTTTCAATTTGAG 968

QY 1136 CAAATTAGAGAGCTTTTGTATCTAAATATATGCTATCCGAGAGCTGATAAGAAATACAA 1195
 DB 969 CAAATTAGAGAGCTTTTGTATCTAAATATATGCTATCCGAGAGCTGATAAGAAATACAA 1028
 QY 1196 GAGCATGAAAGCTGAAAAAGAAATTTGTCAACCCCTCAAGAGCTATAGCTTGTCTAATTC 1255
 DB 1029 GAGCATGAAAGCTGAAAAAGAAATTTGTCAACCCCTCAAGAGCTATAGCTTGTCTAATTC 1088
 QY 1256 AGACTCCATGACATATGAAAGATGTTGAAGCCGTTCAAGAGCTTCAAGATGCTTATCATGA 1315
 DB 1089 AGACTCCATGACATATGAAAGATGTTGAAGCCGTTCAAGAGCTTCAAGATGCTTATCATGA 1148
 QY 1316 AGGCTGCAAGATTTATGAAAGCTGCGACATATGAAAGAGCCCTGCTGAGCTGCGAAGAT 1375
 DB 1149 AGGCTGCAAGATTTATGAAAGCTGCGACATATGAAAGAGCCCTGCTGAGCTGCGAAGAT 1208
 QY 1376 GCTGATGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1435
 DB 1209 GCTGATGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1268
 QY 1436 CATCAAACTAGAAAG 1495
 DB 1269 CATCAAACTAGAAAG 1328
 QY 1496 GGTCTGACTTAAAGCTCCCTGGGCTTCCCATCTTCATGTTGAAAGAGAGAGAGAG 1555
 DB 1329 GGTCTGACTTAAAGCTCCCTGGGCTTCCCATCTTCATGTTGAAAGAGAGAGAGAG 1388
 QY 1556 CCAAGAGTATGTCGAAGAAACTTAAAGTAAACAAT 1598
 DB 1389 CCAAGAGTATGTCGAAGAAACTTAAAGTAAACAAT 1431

RESULT 11
 ADC23488
 ID ADC23488 standard; DNA; 1377 BP.
 AC ADC23488;
 DT 18-DEC-2003 (first entry)
 DE DNA encoding the ligand binding domain of the ERR3 protein.
 KW oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;
 KW ligand binding domain; fertility; birth control; bone remodelling;
 KW cancer; protein coordinate data; gene; ds.
 OS Unidentified.
 FH Key Location/Qualifiers
 FT 1..1377
 FT CDS /*tag= a
 FT /product= "Ligand binding domain of the ERR3 protein"
 XX MO2003064468-A2.
 XX 07-AUG-2003.
 XX 30-JAN-2003; 2003WO-EP000959.
 XX 31-JAN-2002; 2002US-0352551P.
 XX (CNRS) CNRS CENT NAT RECH SCT.
 XX Moras D, Renaud J, Greschik H, Wurtz J;
 XX WPT; 2003-663467/62.
 XX P-PSDB; ADC23486.
 XX New peptide fragment, useful for screening compounds that are agonists or
 XX antagonists of the transcriptional-activating activity of the estrogen-
 XX related receptor 3 (ERR3).
 PT

XX Disclosure; SEQ ID NO 3; 253bp; English.
 PS This invention relates to novel peptide fragments that have an agonistic
 XX or antagonistic effect on the transcriptional-activating activity of the
 CC oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen
 CC related receptor gamma protein. Specifically, the peptide fragment of the
 CC invention comprises a ligand binding domain of ERR3, which becomes
 CC functionally active when fused to a protein containing a DNA binding
 CC domain and mimics the transcriptional-activating activity of the complete
 CC ERR3 protein. As such, the peptide fragment can be used in a screening
 CC method or to design and select compounds that affect ERR3 activity.
 CC Furthermore, the agonists and antagonists of ERR3 are biologically active
 CC compounds that can be used to modulate the oestrogenic response on
 CC fertility, birth control, bone remodelling, breast and prostate cancer.
 CC This polynucleotide sequence is the DNA encoding the ligand binding
 CC pocket of the ERR3 protein of the invention. NOTE: This sequence is given
 CC in the sequence listing but is not further referred to in the
 CC specification.
 XX
 XX Sequence 1377 BP; 375 A; 341 C; 335 G; 326 T; 0 U; 0 Other;
 SQ
 Query Match 46.1%; Score 1377; DB 9; Length 1377;
 Best Local Similarity 100.0%; Pred. No. 2.6e-303;
 Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

126 ATGATTCGGTAGAATTGCTTCTGATCTTTTCCCTGACATGAGGAGAGCTT 185
 1 AAGATTCGGTAGAATTGCTTCTGATCTTTTCCCTGACATGAGGAGAGCTT 60
 186 CTCTGCAAGATTCGAAACAAAGATGACACATTGATTCAGCTGTCCTTCATCAAG 245
 61 CTCTGCAAGATTCGAAACAAAGATGACACATTGATTCAGCTGTCCTTCATCAAG 120
 246 ACGGAACCTTCAGCCCAAGCCCTCCCTGAGGAGACAGGTCACACACAGCCCTGAGC 305
 121 ACGGAACCTTCAGCCCAAGCCCTCCCTGAGGAGACAGGTCACACACAGCCCTGAGC 180
 306 TCTTCAGACGCCAGTGAGAGTACAGTTCAACCATGAATGAGCATGAGACGAGCTTGAC 365
 181 TCTTCAGACGCCAGTGAGAGTACAGTTCAACCATGAATGAGCATGAGACGAGCTTGAC 240
 366 TCGCCACTCTTACCCCTTCTGCTCTCTATCTCTGAGAGTATGAGGCTCTGAGAACTG 425
 241 TCGCCACTCTTACCCCTTCTGCTCTCTATCTCTGAGAGTATGAGGCTCTGAGAACTG 300
 426 TATGATGACTGCTCCAGACCACTTTGTGAAGATCCCCAGACCAAGTGAATCATGCTC 485
 301 TATGATGACTGCTCCAGACCACTTTGTGAAGATCCCCAGACCAAGTGAATCATGCTC 360
 486 AACTGATGATGCCCAAGAGACTGTGTAGTGTGTGTGATCATGCTCTGGGTACCATAT 545
 361 AACTGATGATGCCCAAGAGACTGTGTAGTGTGTGTGATCATGCTCTGGGTACCATAT 420
 546 GGGGTAGATCATGTGAAGCTCTGCAAGCATTTCTTAAGAGACAAATTCAGCAATATA 605
 421 GGGGTAGATCATGTGAAGCTCTGCAAGCATTTCTTAAGAGACAAATTCAGCAATATA 480
 606 GAATACAGCTGCCCGCCAGCAAGATGAATGAATCAAAAGCCGAGACGTAATCCCTGC 665
 481 GAATACAGCTGCCCGCCAGCAAGATGAATGAATCAAAAGCCGAGACGTAATCCCTGC 540
 666 CAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGACATGCTGAAGAAGAGGGGTGCTCTT 725
 541 CAGGCTTGCCGCTTCATGAAGTGTAAAGTGGGACATGCTGAAGAAGAGGGGTGCTCTT 600
 726 GACGAGTACGTGAGAGTCTGGCAGAAATCAAGCCGAGATGATGCGGAGAACAGCCCA 785
 601 GACGAGTACGTGAGAGTCTGGCAGAAATCAAGCCGAGATGATGCGGAGAACAGCCCA 660
 786 TACCTGAACCTCGAGCTGCTCAGCCAGCAAAAAAGCATATATCAAAATGTCTCATAT 845
 661 TACCTGAACCTCGAGCTGCTCAGCCAGCAAAAAAGCATATATCAAAATGTCTCATAT 720

QY 846 TTGTTGTGGCTGTAACCGGAGAAATCTATGCCATGCTGACCTTACCTGCTCCCGACAGT 905
 DB 721 TTGTTGTGGCTGTAACCGGAGAAATCTATGCCATGCTGACCTTACCTGCTCCCGACAGT 780
 QY 906 GACATCAAAAGCCCTCACTACAGTGTGACTTGGCCGACCGAGATGTGGTGTATCATT 965
 DB 781 GACATCAAAAGCCCTCACTACAGTGTGACTTGGCCGACCGAGATGTGGTGTATCATT 840
 QY 966 GGAATGGCGAGACATATTCAGAGCTTCTCCAGCTGTCCCTGGCGAGACGATGAGCTT 1025
 DB 841 GGAATGGCGAGACATATTCAGAGCTTCTCCAGCTGTCCCTGGCGAGACGATGAGCTT 900
 QY 1026 CTGCAGAGCTTGGATGAGTAATTTTGAATCTTGGTGTGCTATACCGGTCTTTGATTT 1085
 DB 901 CTGCAGAGCTTGGATGAGTAATTTTGAATCTTGGTGTGCTATACCGGTCTTTGATTT 960
 QY 1086 GAGATGAACCTGCTATGACAGCATATATATATGAGAGAGAGAGAGAGAGAGAGAG 1145
 DB 961 GAGATGAACCTGCTATGACAGCATATATATATGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 1146 GGCTCTTCTGATCTAATAATATGCTATCTGACAGCTGAGTAAGAAATACAGAGCATGAAG 1205
 DB 1021 GGCTCTTCTGATCTAATAATATGCTATCTGACAGCTGAGTAAGAAATACAGAGCATGAAG 1080
 QY 1206 CTGGAAGAAAGAGATTTTGTCAACCTCMAAGCTATAGCTCTTGTATATTCAGATTCAGT 1265
 DB 1081 CTGGAAGAAAGAGATTTTGTCAACCTCMAAGCTATAGCTCTTGTATATTCAGATTCAGT 1140
 QY 1266 CACATAGAAGATGTGAAGCCCTTCAGAGCTTCAGAGATGTCTTACATGAGAGCTGACAG 1325
 DB 1141 CACATAGAAGATGTGAAGCCCTTCAGAGCTTCAGAGATGTCTTACATGAGAGCTGACAG 1200
 QY 1326 GATTATGAAGCTGCGCAGACATGAGAGACCTTGTGAGCTGCGCAAGATGCTGATGACA 1385
 DB 1201 GATTATGAAGCTGCGCAGACATGAGAGACCTTGTGAGCTGCGCAAGATGCTGATGACA 1260
 QY 1386 CTGCACTCTCGAGGACAGCCCTTACCAAGGCGCGAGACATTTCTACAACTCAAACTTA 1445
 DB 1261 CTGCACTCTCGAGGACAGCCCTTACCAAGGCGCGAGACATTTCTACAACTCAAACTTA 1320
 QY 1446 GAAGGCAAAAGTCCCAATGCACAAACTTTTGTGAATGTGAGAGCCCAAGGCTGGA 1502
 DB 1321 GAAGGCAAAAGTCCCAATGCACAAACTTTTGTGAATGTGAGAGCCCAAGGCTGGA 1377

RESULT 12
 AAA40078
 ID AAA40078 standard; cDNA; 1377 BP.
 XX
 AC AAA40078;
 XX
 DT 19-OCT-2000 (first entry)
 XX
 DE Human ERRgamma coding region cDNA.
 XX
 KW Human; ERRgamma; brain; estrogen-related receptor gamma; cytosolic;
 KW lipid metabolism; fatty acid synthesis; antiarteriosclerotic; treatment;
 KW drug development; diabetes; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200026365-A1.
 XX
 PD 11-MAY-2000.
 XX
 PF 02-NOV-1999; 99WO-JP006097.
 XX
 PR 04-NOV-1998; 98JP-00313194.
 XX
 PA (KAZU-) KAZUSA DNA RES INST FOUND.
 PA (TAIS) TAISHO PHARM CO LTD.
 XX

PI Ohara O, Nagase T, Nomura N, Takayama K, Toyoda H, Yoshimoto M;
 XX WPI; 2000-365614/31.
 DR P-PSDB; AAB09965.
 XX
 PT An estrogen-related receptor gamma protein with lipid metabolism
 XX regulatory and fatty acid synthesis functions, and its encoding gene,
 PT useful as drugs and in developing drugs for treatment of e.g.
 XX arteriosclerosis and diabetes.
 PS Claim 2a; Page 24-26; 38pp; Japanese.
 XX
 CC This invention describes a novel estrogen-related receptor gamma
 CC (ERGamma) protein (1) and its variants maintaining receptor activity,
 CC having lipid metabolism regulatory and fatty acid synthesis functions.
 CC The product of the invention has cytosolic and antiarteriosclerotic
 CC activity. The protein and gene are useful as drugs and in developing
 CC drugs for treatment of e.g. arteriosclerosis and diabetes. The protein is
 CC obtained by cloning human brain-originated cDNA library. This sequence
 CC encodes the human brain ERGamma protein which is described in the method
 CC of the invention
 XX
 SQ Sequence 1377 BP; 374 A; 341 C; 336 G; 326 T; 0 U; 0 Other;
 Query Match 46.1%; Score 1375.4; DB 3; Length 1377;
 Best Local Similarity 99.9%; Pred. No. 5.9e-303;
 Matches 1376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 126 ATGATTCGGTGAAGACTTGGCTTCTGATCTTTTCCCTGCACTACGAGGAAGCTT 185
 DB 1 ATGATTCGGTGAAGACTTGGCTTCTGATCTTTTCCCTGCACTACGAGGAAGCTT 60
 QY 186 CTGCGAGATGTCAAAACAAGATCGACATGATTCAGCTGTGTGCTTCATCAAG 245
 DB 61 CTGCGAGATGTCAAAACAAGATCGACATGATTCAGCTGTGTGCTTCATCAAG 120
 QY 246 ACGGAACTTCGACGCCAGCTCCCTGAGGAGACAGCTCAACCAACAGCCTGTGGTGC 305
 DB 121 ACGGAACTTCGACGCCAGCTCCCTGAGGAGACAGCTCAACCAACAGCCTGTGGTGC 180
 QY 306 TCTTCAGACGCGATGGGAGTACATGTTCAACATGATGCGCTCAAGACGAGCTTGAAC 365
 DB 181 TCTTCAGACGCGATGGGAGTACATGTTCAACATGATGCGCTCAAGACGAGCTTGAAC 240
 QY 366 TCGGCACCTCTCTACCTTCTGCTCTATCTGAGGAGTGTGTGGGCTGTCAAGAACTG 425
 DB 241 TCGGCACCTCTCTACCTTCTGCTCTATCTGAGGAGTGTGTGGGCTGTCAAGAACTG 300
 QY 426 TATGATGACTGCTCCAGACACCATTTGTAAGATCCCAACCAAGTGTGAATACATGCTC 485
 DB 301 TATGATGACTGCTCCAGACACCATTTGTAAGATCCCAACCAAGTGTGAATACATGCTC 360
 QY 486 AACTCGATGCCAGAGACATGCTGTTAGTGTGTGTGATCAATGCTTCGAGTACACTAT 545
 DB 361 AACTCGATGCCAGAGACATGCTGTTAGTGTGTGTGATCAATGCTTCGAGTACACTAT 420
 QY 546 GGGGTAGCATCATGTGAAGCTGCAAGGCAATTTCTCAAGAGACATTTCAAGCAATATA 605
 DB 421 GGGGTAGCATCATGTGAAGCTGCAAGGCAATTTCTCAAGAGACATTTCAAGCAATATA 480
 QY 606 GAATACAGCTGCTCGCCAGATGATGTAATACAAAGCGCAAGCTGAATCTCTGC 665
 DB 481 GAATACAGCTGCTCGCCAGATGATGTAATACAAAGCGCAAGCTGAATCTCTGC 540
 QY 666 CAGGCTTGGCGCTTCATGAAGTGTAAAGTGGGCAATGCTGAAGAAAGGGGTGCGTCTT 725
 DB 541 CAGGCTTGGCGCTTCATGAAGTGTAAAGTGGGCAATGCTGAAGAAAGGGGTGCGTCTT 600
 QY 726 GACAGAGTACGTGAGGTGCGCAGAGTACAAAGCGCAGAGATGATGCGAGAAACAGCCCA 785
 DB 601 GACAGAGTACGTGAGGTGCGCAGAGTACAAAGCGCAGAGATGATGCGAGAAACAGCCCA 660
 QY 786 TACCTGAACCTCTCAGCTGGTTGACGACAGCAAAAAGCCATATTAACAGATTGTCTCAT 845

DB 661 TACCTGAACCTCTCAGCTGGTTGACGACAGCAAAAAGCCATATTAACAGATTGTCTCAT 720
 QY 846 TTGTTGGTGGCTGAACCCGAGAAAGATCTATGCAATGCTGACCTTACTGCCCCGACAGT 905
 DB 721 TTGTTGGTGGCTGAACCCGAGAAAGATCTATGCAATGCTGACCTTACTGCCCCGACAGT 780
 QY 906 GACATCAAGCCCTCACTACCTGTGTGACTTGGCCGACCGAGAGTGGTGTATCATTT 965
 DB 781 GACATCAAGCCCTCACTACCTGTGTGACTTGGCCGACCGAGAGTGGTGTATCATTT 840
 QY 966 GGATGGGCGAAGATATTCAGGCTTCCAGAGCTGCTCCGCGGAGACAGATGAGCCTT 1025
 DB 841 GGATGGGCGAAGATATTCAGGCTTCCAGAGCTGCTCCGCGGAGACAGATGAGCCTT 900
 QY 1026 CTGCAAGTGTGATGAGAAATTTGATCCTTGTGTGTGTATACCGGTCTCTTCAATT 1085
 DB 901 CTGCAAGTGTGATGAGAAATTTGATCCTTGTGTGTGTATACCGGTCTCTTGTGTT 960
 QY 1086 GAGGATGAACCTGTCTATGACAGATATATATATGAGACGACATCAATTAATGCA 1145
 DB 961 GAGGATGAACCTGTCTATGACAGATATATATATGAGACGACATCAATTAATGCA 1020
 QY 1146 GGCCTTCTGATCTAATATATATGCTATCCGACGCTGTAAGAAATACAGAGCATGAG 1205
 DB 1021 GGCCTTCTGATCTAATATATATGCTATCCGACGCTGTAAGAAATACAGAGCATGAG 1080
 QY 1206 CTGGAAGAAAGAAATTTGCTACCTCTCAAGCTATAGCTTGTCTAATTCAGATCCATG 1265
 DB 1081 CTGGAAGAAAGAAATTTGCTACCTCTCAAGCTATAGCTTGTCTAATTCAGATCCATG 1140
 QY 1266 CACATGAGATGTTGAAGCTGTTCAAGAGCTTCAGAGTGTCTTACATGAAGGCTTCAG 1325
 DB 1141 CACATGAGATGTTGAAGCTGTTCAAGAGCTTCAGAGTGTCTTACATGAAGGCTTCAG 1200
 QY 1326 GATTATGAAGCTGGCGACGACATGAGAAACCTGTGTGAGCTGGCAAGATGCTGATGACA 1385
 DB 1201 GATTATGAAGCTGGCGACGACATGAGAAACCTGTGTGAGCTGGCAAGATGCTGATGACA 1260
 QY 1386 CTGCCACTCCTGAGGAGACCTCTACCAAGCCGTGAGCATTTCTAACAATCAACTA 1445
 DB 1261 CTGCCACTCCTGAGGAGACCTCTACCAAGCCGTGAGCATTTCTAACAATCAACTA 1320
 QY 1446 GAAAGCAAGTCCCAATGACAAACTTTTGTGAAATGTTGAGAGCAAGGCTTGA 1502
 DB 1321 GAAAGCAAGTCCCAATGACAAACTTTTGTGAAATGTTGAGAGCAAGGCTTGA 1377
 RESULT 13
 AAA72605
 ID AAA72605 standard; cDNA; 1996 BP.
 XX
 AC AAA72605;
 XX
 DT 27-NOV-2000 (first entry)
 XX
 DE Human oestrogen related receptor 4 (ERR4) cDNA sequence.
 XX
 KW Oestrogen related receptor 4; ERR4; antiinflammatory; cytosolic;
 KW antidiabetic; nephrotropic; osteopathic; inflammation; cancer;
 KW osteoporosis; diabetes; kidney disease; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 140..1441
 FT CDS /tag= a
 FT /product= "ERR4"
 FT /note= "Oestrogen related receptor 4"
 XX
 PN MO200042180-A1.
 XX
 PD 20-JUL-2000.


```

XX 27 AUG-1997; 97US-00570902.
PR 21-OCT-1997; 97US-0062902P.
PR 19-MAR-1998; 98US-0078633P.
XX
XX (MERI) MERCK & CO INC.
XX
XX Chen F;
XX
XX WPI: 1999-190586/16.
XX P-PSDB; AAY03836.
XX
XX Newly purified DNA polynucleotides encoding human nuclear trans-acting
XX receptor proteins - useful in the diagnosis, treatment and prophylaxis of
XX cell differentiation, development and physiological function.
XX
XX Claim 9; Page 41-42; 82pp; English.
XX
XX The invention relates to DNA molecules encoding human nuclear receptor
XX (NMR) proteins nmr1 and nmr2. The nmr proteins (including mutants and/or
XX fragments) form pharmaceutical compositions that are useful in the
XX diagnosis, treatment and prophylaxis of cell differentiation, development
XX and physiological function. The proteins are also useful for identifying
XX downstream target genes and ligands regulating their activity. In
XX particular, fusion constructs (especially glutathione S-transferase (GST)
XX -nmr1 and/or GST-nmr2) expressing fusion proteins are useful in screening
XX for (ant)agonists that are useful as modulators in cell differentiation,
XX development and physiological function. Antibodies for the nmr proteins
XX are useful for measuring the levels of these proteins. The present
XX sequence represents a DNA encoding a human nmr1 protein.
XX
XX Sequence 2807 BP; 701 A; 777 C; 770 G; 559 T; 0 U; 0 Other;
XX
XX Query Match 24.0%; Score 715; DB 2; Length 2807;
XX Best Local Similarity 71.5%; Pred. No. 1.8e-152;
XX Matches 956; Conservative 0; Mismatches 375; Indels 6; Gaps 1;
XX
XX 179 AAGAGCTTCTCTGAGATGTCAAACAAAGATGACACATTGATTCAGCTGTTCTGCTT 238
XX |||||
XX 934 AAGGCTGTGAACAGAGATGCTCCGACACAGGCACTGGGCTCAGCTGCGGCTCTCT 993
XX
XX 239 CATCAAGACGGAACCTTCCAGGCCACGCTCCCTGAGGAGACAGGTCAACACACAGGCC 298
XX |||||
XX 994 CATCAAGATGAGCGCTTCCAGCCGCTCTCGGACATGATGCTTCCAGCCACACAGGCC 1053
XX
XX 299 TGGTGGCTCTTGAAGCGCAGTGGAGCTACAGTTCAACCATGATGAGCCATCAAGACG 358
XX |||||
XX 1054 CAGTGGCTGTGCGGAGCGCAGCGGCGCTTGGCCCTGCGCCCTGGGACCCACCG 1113
XX
XX 359 ACTTGACTGGCACTCTCTACCTTCTGCTCTATCTGCTGGAGGTAGTGGGCTCTGAG 418
XX |||||
XX 1114 TCTGGAGCTGCGCAC-----CATGTTTGAAGCGCGCGGCTGGAGGACCCCATGCGG 1167
XX
XX 419 GAAACGTGATGATGATGCTCCAGGACCATTTGTTGATGATGATGATGATGATGATGAT 478
XX |||||
XX 1168 CAAAGAGCTGAGAGACTGTGCGAGCGCATGAGAGACTGCGCATCAAGTGAAGTGA 1227
XX
XX 479 CATGCTCAACTGATGCTCCAGAGAGACTGATGATGATGATGATGATGATGATGATGAT 538
XX |||||
XX 1228 CATGCTCAAGCGCATCCCAAGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1287
XX
XX 539 CCACTATGGGGTAGCATGTGTAAGCTTGAAGGCAATTTCTTAAGAGGCAATTTCAAG 598
XX |||||
XX 1288 CCACTATGGGGTAGCATGTGTAAGCTTGAAGGCAATTTCTTAAGAGGCAATTTCAAG 1347
XX
XX 599 CAAATATAGATATAGAGCGCTCCGACGATGATGATGATGATGATGATGATGATGATGAT 658
XX |||||
XX 1348 GAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1407
XX |||||
XX 659 ATCTGCTCAGGCTTGGCGCTTCAAGAGTGTAAAAGTGGGATGCTGTAAGAGGAGTGT 718
XX |||||
XX 1408 GTCTGCTCAGGCTTGGCGCTTCAAGAGTGTAAAAGTGGGATGCTGTAAGAGGAGTGT 1467
XX

```

```

QY 719 GGTCTTTGACAGAGTACGTGAGGCTGCGACAGATGACAAAGCCGAGATGATGCGGAGAA 778
DB 1468 GGGCTTTGATGAGTGGCTGGAGGCGCTGACAGAAATCAAGCCAGCTGAGACTCAAGAG 1527
QY 779 CAGCCCATACCTGAGACCTCAGCTGATTCAGCGCAGCCAAAAGCATATACAGATTTGT 838
DB 1528 CAGCCCATACCTGAGACTTACAAATTTCTCCACTCTGTAAAGCATTTGACCAAGATTTGT 1587
QY 839 CTACATTTGTTGTGCTGTAACCGAGAAATCTATGTCATGCTGACCTTACTGTGCC 898
DB 1588 CTCATACCTACTGTGTGCTGAGCCGAGCAAGCTCATGTCATGCTGCTGCTGCTGCTGCTG 1647
QY 899 CGACGTGATCATCAAAAGCCCTCACTACACTGTGTGATCTTGGCCGACCGAGATTTGGT 958
DB 1648 TGAGGGGAGCATCAAGGCGCTTACCACTCTGTGATCCTGGAGACCGAGAGTGTGTGT 1707
QY 959 TATCATTTGATGGGGAGGAGCATATTCAGGCTTTCACGCTGTCCCTGGCGGACAGAT 1018
DB 1708 CATCATTTGCTGGGCGACAGACATCCAGGCTTTCAGGCTTCTCCCTGGGGAGCCAGAT 1767
QY 1019 GAGCCTTTCGAGAGTGTGATGAGAAATTTGATTCCTTGTGTGTATACCGATCTCT 1078
DB 1768 GAGCCTTTCGAGAGTGTGATGAGAAATTTGATTCCTTGTGTGTATACCGATCTCT 1827
QY 1079 TTCAATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
DB 1828 GCGCTTACGACGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1887
QY 1139 ATTGAGAGGCTTCTTGTATCTTAATATGCTATCTCCGAGCTGTGTGTGTGTGTGTGTGT 1198
DB 1888 CCTCGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1947
QY 1199 CATGAGCTGAAAAAGAAATTTGTACCTTCAAGGCTATGCTCTTGTATTTGAG 1258
DB 1948 GCTCAAGGTGAGAGAGAGAGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2007
QY 1259 CTCATGACCATGAGAGATGTTGAAGCCTTCAAGGCTTGTGTGTGTGTGTGTGTGTGTGT 1318
DB 2008 TTCCATGTCATGAGAGATCTAGAGCTGTCCAGAGCTGTGACAGGACCTGTGTGACAGAG 2067
QY 1319 GCTGAGAGATTTGAGAGCTGTGAGAGATGAGAGAGCTGTGAGAGCTGTGAGAGAGTGT 1378
DB 2068 ACTGACAGAGCTACGAGCTAGAGCGCATGAGAGAGCTGTGAGAGAGCTGTGAGAGAGTGT 2127
QY 1379 GATGACATGCACTCTGAGGAGAGCTGTACCAAGGCTGTGAGAGAGCTGTGAGAGAGTGT 1438
DB 2128 GCTGACATGCGGCTGTGAGGAGAGCGCGCAAGGCTGTGAGAGAGCTGTGAGAGAGTGT 2187
QY 1439 CAACTGAGAGGCAAGTCCCATGACAAATTTTTTGGAAATGTTGAGAGGCAAGGT 1498
DB 2188 CAAACTGAGAGGCAAGTCCCATGACAAATTTTTTGGAAATGTTGAGAGGCAAGGT 2247
QY 1499 CTGACTTAAAGCTCCCT 1515
DB 2248 CTGGGCGAGGCTGACT 2264

```

RESULT 15
AAA09800
ID AAA09800 standard; DNA; 2807 BP.

XX AAA09800;
XX
XX 05-JUL-2000 (first entry)
XX
XX Human nuclear receptor nmr1 gene.
XX
XX Human nuclear receptor protein-1; nmr1; physiological function; ss;
XX cell development and differentiation controller; gene expression.
XX Homo sapiens.
XX
XX US6054295-A.

```

XX 25-APR-2000.
XX
XX 26-AUG-1998; 98US-00141000.
XX
XX 27-AUG-1997; 97US-0057090P.
XX 21-OCT-1997; 97US-0062922P.
XX 19-MAR-1998; 98US-0078633P.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Chen F;
XX
XX WPI: 2000-328352/28.
XX P-PSDB; AAY83822.
XX
XX New polynucleotide encoding human nuclear receptor protein-1 is useful
XX for screening compounds that acts as modulators of cell differentiation,
XX cell development and physiological function.
XX
XX Claim 9; Col 57-58; 58pp; English.
XX
XX This sequence represents the DNA encoding human nuclear receptor protein-
XX 1 (NMN1). The protein is a cell development and differentiation
XX controller which is useful for screening compounds that act as modulators
XX of cell differentiation, cell development and physiological functions or
XX for treating and diagnosing disorders associated with altered expression
XX of the DNA
XX
XX Sequence 2807 BP; 701 A; 777 C; 770 G; 559 T; 0 U; 0 Other;
XX
XX Query Match 24.0%; Score 715; DB 3; Length 2807;
XX Best Local Similarity 71.5%; Pred. No. 1.8e-152;
XX Matches 956; Conservative 0; Mismatches 375; Indels 6; Gaps 1;
XX
XX 179 AGAGCTTCTGCGAATGATGTAACAAAGATGACACATTGATTCAGCTGTGCTCTT 238
XX 934 AGGCTGTGTGAACAGAGATGCTCTGAGACGACAGGACCTGGGCTCCAGCTCGGCTCCTT 993
XX
XX 239 CATCAAGCGGAGACCTTCAGGCGCAGCTCCTGACGACGCGTCAACCCACCAAGGCC 298
XX 994 CATCAAGCTGAGCGCTGACGCGCTCCTCGGCGCATGATGCTCCAGCCACCAAGGCC 1053
XX
XX 299 TGGTGGCTCTTCAGACGCGCAGTGGAGCTACAGTTCAACATGATGGCCATCGAAAG 358
XX 1054 CAGTGGCTCGTCCGACGCGCAGGCGGCTTGGCGCTGGCCCTGGGACCCACGCGCAAG 1113
XX
XX 359 ACTTGACTCGCAGCTCTTACCTTCTGCTCTATCTGAGAGGTAGTGGGCTGTGAG 418
XX 1114 TCTTGACTCGCAGC-----CATGTTGACAGGCGCGGCTGGGAGGCACTCCATGCG 1167
XX
XX 419 GAACCTGATGATGATGCTGCTGACAGACATGTTGAAAGATCCCGACCAAGTGTGAATA 478
XX 1168 CAAAGCTGAGAGACTGTGCTGACAGGCGCATATGAGAGACTCGGCATTAAGTGGAGTA 1227
XX
XX 479 CATGCTCAACTCGATGCGCCAGAGAGACTGTGTTTATGTTGTTGATGATCGCTTGGGTA 538
XX 1228 CATGCTCAAGCGCCATCCCAAGCGCGTGTGCTGTGCGGGAGCAATGCTCTGGCTA 1287
XX
XX 539 CCATCTAAGGGGTAGCATCATGATGAAAGCTGCAAGGCATTTCTCAAGAGGACATTCAG 598
XX 1288 CCAGCTACGGCGTGGCTCTGCGAGGCTTGGCAAGGCTTTCTTCAAGAGGACTATCCAA 1347
XX
XX 599 CAATATAGATATACAGCTGCGCTGCGACAGATGATGTAATCAAAAGCGCAGACGTAA 658
XX 1348 GAAATATTGAGTACAGCTGCGCGCCGCGACCAAGAGTGCAGATCAACCAAGCGGCGCA 1407
XX
XX 659 ATCTGCGAGGCTTGCCTTCATGAAAGTGTAAAGTGGGCAATGCTGAAAGAGGAGT 718
XX 1408 GTCTGCGAGGCTGCGGCTTCATGAATGCTCAAGTGGGATGCTGAAGAGAGGT 1467
XX
XX 719 GCGCTTGAAGAGTATGAGAGGTGCGAGAGATCAAGCGCAGAGTATGATGCGAGAA 778
XX

```

```

Db 1468 GCGCTTATCATGAGTGGGTGAGGCGCTCAGAAAATACAAAGCGAGCGCTGACTGAGAG 1527
XX
XX 779 CAGCCCATACCTGTAACCTCTAGCTGTTCAGCCGACCCAAAAGCATATTAAGATTTGT 838
XX 1528 CAGCCCATACCTGAGCTTACAAATTTCTCCACCTGTAAAAAAGCCATTAAGCAAGATTGT 1587
XX
XX 839 CTCACATTTGTTGGGTGAGTGAACCGGAGAAAGATCATAGCCATGAGCTTACCTGATCC 898
XX 1588 CTCATACCTACTGAGTGGGTGAGCCGAGACAGCTCTAGTCCATGCTCTCCCTGATATGCC 1647
XX
XX 899 CGACAGTACATCAAAAGCCTCACTACACTGTGTGACTTGGCCGACCGAGAGTTGTGTT 958
XX 1648 TGAGGGGAGACATCAAGGCGCTGACCACTCTGTGACTGTGACAGACGAGAGCTTGTGT 1707
XX
XX 959 TATCATTTGATGGGCGGAAGCATTTCCAGGCTTCTCCAGCTGTTCCTCGGCGGACCAAT 1018
XX 1708 CATCATTTGGCTGGGCGCAAGCATCCAGGCTTCTCAAGGCTCTCCCTGGGCGGACCAAT 1767
XX
XX 1019 GAGCCTTCTGCGAGAGTGTGATGAGAAATTTGATCCTGTGTGTGATATACCGGCTCT 1078
XX 1768 GAGCCTGCTGAGAGTGTGCTGATGAGAAATCTCATCTTGGGCGATCGTGTACCGCTGCT 1827
XX
XX 1079 TTCAATTGAGGATGAACCTTGTCTATGACAGCATTAATTAATGAGACGAGACCAAGTCAA 1138
XX 1828 GCGCTTACGACGACAAAGCTGTGTGTGAGCTGAGAGCTACATCATGATGAGAGCACTCCG 1887
XX
XX 1139 ATTGACGAGGCTTGTATCTAATTAATGCTATCTGTAGCTGTGTAAGAAATTAATCAAG 1198
XX 1888 CTTGCGGGGCTGTGAGCTCTACCGGCGCATCTGTGAGCTGTGACGAGTACAAAGAA 1947
XX
XX 1199 CMTGAGCTGGAAGAAAGAGAAATTTGTACCCCTCAAGCTTATAGCTTGTAAATTCAGA 1258
XX 1948 GCTCAAGGTGAGAGAGAGAGGTTTGTAGCGCTCAAGGCCCTGTGCGCACTCCGA 2007
XX
XX 1259 CTCATGCAATGAAGATGTTGAAGCGCTTCAGAGCTTCAGAGTGTCTTAATGAAGC 1318
XX 2008 TTCCATGATACATCAGAGATCTAGAGGCTGTGTCAGAAAGCTGCAAGACCTGTGCAAGAGC 2067
XX
XX 1319 GCTGACAGATTATGAAGCTGSCGACCATGAGAGACCTGTGAGCTGTGCAAGATGCT 1378
XX 2068 ACTGACAGACTACAGACTGAGCGCAGCGCATGAGAGAGCTTGGAGAGAGCGGCAAGCTCT 2127
XX
XX 1379 GATGACACTGCACTCTGAGGCGAGACCTTACCAAGCGCGTGAAGCATTTCTCAACAT 1438
XX 2128 GCTACACTGCGCGTGTGCTGCGGAGAGCGCGCGCAAGCGCTGTGACACTTCTATAGGT 2187
XX
XX 1439 CAAACTGAAAGGCAAGTCCCATATGCAAACTTTTGTGAATGTTGAGGCGCAAGGT 1498
XX 2188 CAAACTGCAAGGCAAAAGTGGCCATGCAAACTTCTGTGAGATGCTGAGAGCGCAAGGC 2247
XX
XX 1499 CTGACTAAAGCTCCCT 1515
XX 2248 CTGGGCGAGGCTGACT 2264
XX

```

Search completed: August 9, 2004, 23:43:00
 Job time : 1138 secs

REFERENCE
AUTHORS
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunaka, T., Tachibana, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384 format
sequencing pipeline with 384 multiplexed capillary sequencer
JOURNAL
MEDLINE
20030913
PUBMED
11076861

REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
JOURNAL
NATURE
409, 685-690 (2001)

REFERENCE
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
NATURE
420, 563-573 (2002)

JOURNAL
REFERENCE
AUTHORS
6
(Bases 1 to 4835)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Katsukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohmoto, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, S., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
Direct Submission

JOURNAL
TITLE
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsukuba-shi, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES
SOURCE
1. 4835
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D630039B07"
/db_xref="MG1:2432648"
/db_xref="taxon:10090"
/clone="D630039B07"
/tissue type="kidney"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
281. 1588
/note="unmated protein product; estrogen-related receptor gamma (MG1 MG1:1347056, GB|NM_011935, evidence: BLASTV, 99%, match=1444)
putative"
/codon_start=1
/protein_id="BAC35120.1"
/db_xref="GI:26342927"

ORIGIN
Query Match 64.1%; Score 1913.8; DB 11; Length 4835;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 2353; Conservative 0; Mismatches 412; Indels 52; Gaps 14;
/translation="MSNRDRHIDSCSFRITKREPSPAELTDSVNHHSFGSSDASGS
YSTINQHNQNDSPPLPSAPILIGSGSVRLYDDCSSTIVEDPTKCEYNLSMPK
RLCIVCGDIASGYHGVASCEACKAFKRTIGANTIBYSPANBEKTRBRKSCQAC
RPMKCIKYOMLEKGVRLDRVGRGKRRIDANSYPYNPOLVOPAKYIVSHL
LVAEPEKLYAMDVPDSDIKATLTLCDLAPRELVIITGMKHI PGFSTLSAOMS
LLOSAMEKILGVVYRSISFEDLEIVADDTYMDROSLKACLDLNNH ILDTLVKVK
SMKEKEFEVTLTKALAINSDSMHIEDVENOKLDVTHRALQDYBAGOHMEDPFRAG
KMLMTLPILRQTSIRAVQHFVNIKLEKVPWKKLFLEMLKAV"

163 CCTTGCACTACGAGGAAGAGCTTCTCTGCAAGATGTCAACAAAGATGACACATTGATT 222
249 CACTGTTCAGATTGGAAAGGCTTCTCTGCAAGATGTCAACAAAGATGACACATTGATT 308
223 CCAGCTGTTGCTCTTCATCAAGACGAACTTCAGCCGACGCTCCCTGAGGACGACG 282
309 CCAGCTGTTGCTCTTCATCAAGACGAACTTCAGCCGACGCTCCCTGAGGACGACG 368
283 TCAACCAACCAAGCCCTGAGGCTCTTCAGACGCAAGTGGAGCTAACATTCAACATGA 342
369 TCAACCAACCAAGCCCTGAGGCTCTTCAGACGCAAGTGGAGCTAACATTCAACATGA 428
343 ATGGCATCAAGAACGGAATTGATCTGCACTCTTACCTTTCTGCTCTATCTGGAG 402
429 ATGGCATCAAGAACGGAATTGATCTGCACTCTTACCTTTCTGCTCTATCTGGAG 488
403 GTAGGGGCTCTGAGAAACGTATGATGATGCTGCTCCAGACCACTGTTGAAGATCCCC 462
489 GCAAGGGGCTCTGAGAAACGTATGATGATGCTGCTCCAGACCACTGTTGAAGATCCCC 548
463 AGACCAAGTGAATATCATGTCTCACTGATGATGATGATGATGATGATGATGATGATG 522
549 AGACCAAGTGAATATCATGTCTCACTGATGATGATGATGATGATGATGATGATGATG 608
523 ACATGCTCTTGGGATACCATATGATGATGATGATGATGATGATGATGATGATGATGAT 582
609 ACATGCTCTTGGGATACCATATGATGATGATGATGATGATGATGATGATGATGATGAT 668
583 AGAGACCAATCAAGGCAATATATGAAATACAGTGGCTGCTCCAGACCAAGTGAAGTGA 642
669 AGAGACCAATCAAGGCAATATATGAAATACAGTGGCTGCTCCAGACCAAGTGAAGTGA 728
643 CAAAGCGGACGTAATCTCTGCAAGGCTTCCGCTTCATGAAGTGTAAAGTGGGCA 702
729 CAAAGCGGACGTAATCTCTGCAAGGCTTCCGCTTCATGAAGTGTAAAGTGGGCA 788
703 TGTGGAAGAAAGGGGTGGTGTGTTGACAGAGTACGTGAGGTGCGGCAAGATGACGCA 762
789 TGTGGAAGAAAGGGGTGGTGTGTTGACAGAGTACGTGAGGTGCGGCAAGATGACGCA 848
763 GGATAGATGCGGAGAAAGAGCCATACCTGAACTGAGCTGATGACGAGGCAAAAGG 822
849 GAATAGATGCGGAGAAAGAGCCATACCTGAACTGAGCTGATGACGAGGCAAAAGG 908
823 CATTAACCAAGATTGTCTACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 882
909 CATTAACCAAGATTGTCTACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 968
883 CTGACCTTACTGTCCTCCGACAGTACATCAAGGCTTCATCACTGTGTGACTTTGGCGG 942
969 CTGACCTTACTGTCCTCCGACAGTACATCAAGGCTTCATCACTGTGTGACTTTGGCGG 1028
943 ACCGAGAGTGTGTGTATCATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1002
1029 ACCGAGAGTGTGTGTATCATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1088
1003 CCTGTGCGGACCAAGATGAGCTTCTGCAAGTGTGTTGATGATGATGATGATGATGATG 1062
1089 CCTGTGCGGACCAAGATGAGCTTCTGCAAGTGTGATGATGATGATGATGATGATGATG 1148

QY 1063 TCGTATACCGGTCTCTTTCATTTGAGATGAACCTGTATGACAGACGATTTATATATGG 1122
 Db 1149 TTGTGACCGATGCGTTTGGTGGATGAACTGTCTATGACAGACGATTTATATATGG 1208
 QY 1123 AGGAGACGATGCCAAATTAAGAGCGCTTCTGTATTAATATATATGCTATCTGACGCTGG 1182
 Db 1209 ATGAGACGATGCCAAATTAAGAGCGCTTCTGTATTAATATATGCTATCTGACGCTGG 1268
 QY 1183 TAAAGAAATACAGAGACGATGAAGCGTGGAAAAAGAAATTTGTACCCCTCAAGCATATG 1242
 Db 1269 TGAAGAGATACAGAGACGATGAAGCGTGGAAAAAGAAATTTGTACCCCTCAAGCATATG 1328
 QY 1243 CTTCTGCTATTTACAGATGCTCATGACATAGAGATTTGAAAGCGCTTCAAGAGCTTCAG 1302
 Db 1329 CTCTGCTATTTACAGATGCTCATGACATAGAGATTTGAAAGCGCTTCAAGAGCTTCAG 1388
 QY 1303 ATGTCTTACATGAAGCGCTGACAGGATTTAGAGCTGGCCAGACATGAAAGACCTCTGTC 1362
 Db 1389 ATGTCTTACATGAAGCGCTGACAGGATTTAGAGCTGGCCAGACATGAAAGACCTCTGTC 1448
 QY 1363 GAGCTGGCAAGATGCTGATGACCTGCACTCTGAGGAGACCTTACCAAGGCGCTGG 1422
 Db 1449 GTGACAGCAGATGCTGATGACCTGCACTCTGAGGAGACCTTACCAAGGCGCTGG 1508
 QY 1423 AGCATTTCTACAACTCAATCAAACTAGAAAGCAAACTGCCAATGACAAACTTTTGGAAA 1482
 Db 1509 AGCATTTCTACAACTCAATCAAACTAGAAAGCAAACTGCCAATGACAAACTTTTGGAAA 1568
 QY 1483 TGTGGAGGCGAAGGTCTGACTTAAAG--CTCCCTGGGCTCTCCCATCTTCAATGTTGA 1540
 Db 1569 TGTGGAGGCGAAGGTCTGACTTAAAG--CTCCCTGGGCTCTCCCATCTTCAATGTTGA 1627
 QY 1541 AAGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1600
 Db 1628 AAGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1687
 QY 1601 AAATCAACAGACTGCACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1660
 Db 1688 AAATCAACAGACTGCACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1747
 QY 1661 TGCATAGGCTCTGATGAGTCTTCTTCACTTTCT--CATCATCTTCTTCTCTTCTTCT 1719
 Db 1748 GCTACATCTCTGATGAGTCTTCTTCACTTTCT--CATCATCTTCTTCTCTTCTTCT 1807
 QY 1720 CCAATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1779
 Db 1808 CATTCTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1866
 QY 1780 GCTTCTTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1839
 Db 1867 -----TTTCCCTCTCCCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1921
 QY 1840 CTTTCTTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1899
 Db 1922 CTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1981
 QY 1900 AAATAGCTTATGTTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1956
 Db 1982 TCACATCTCTCTCCCACTCTGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2041
 QY 1957 TCTTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2013
 Db 2042 TTACTCTTGAATCTCTTCT 2101
 QY 2014 TTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2063
 Db 2102 TTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2161
 QY 2064 --AAGAGAGATGAGAGCGAGCGCGCGCAAGAGATGAGATGAGATGAGATGAGATGAGATG 2121
 Db 2162 GAGAGAGATGAGAGCGAGCGCGCGCGCAAGAGATGAGATGAGATGAGATGAGATGAGATG 2221

QY 2122 AACTTATGTGACCACTACCGTCCCAATGACTAAGSAAATCAAGAGAGAACCAAGCT 2181
 Db 2222 AACTTATGTGACCACTACCGTCCCAATGACTAAGSAAATCAAGAGAGAACCAAGCT 2275
 QY 2182 TCCATAAAGTACAGTCAACATATACAAATGACTAGAGTCAAGTATTAATTTCAATGGA 2241
 Db 2276 TACCTAAGTACAGTCAACATATACAAATGACTAGAGTCAAGTATTAATTTCAATGGA 2334
 QY 2242 GCAGCTCTTAATTAAGCAATTAAGCAAGTTCATGCGCTGCTTCTTATGATGCTTTT 2301
 Db 2335 GCAGCTCTTAATTAAGCAATTAAGCAAGTTCATGCGCTGCTTCTTATGATGCTTTT 2389
 QY 2302 CCATCTAGATGAGTACAGCAATTTGATGCTTATATGTTTTTCAAGCTTCCAGGAT 2361
 Db 2390 TCATCTAGATGAGTACAGCAATTTGATGCTTATATGTTTTTCAAGCTTCCAGGAT 2449
 QY 2362 TTGTTAGTTAGTACATATGATGCTTTTCAAGGAAATGATTTAATGCTTTATCATG 2421
 Db 2450 TGTGTTAGTTAGTACATATGATGCTTTTCAAGGAAATGATTTAATGCTTTATCATG 2509
 QY 2422 CAATCTAAGAGAAATTAAGAAATCTGCAATTTTGTCTGGCTTTGAACAATTAAGACA 2481
 Db 2510 CAATCTAAGAGAGGAGTAAAGAAATCCGCAACCTGCTGGCTTTGAACAATTT--GAACA 2567
 QY 2482 ATATTAAGAGCAAAATTAATCTGAGAGAAATTTTAAAAATGTTTTGTTCTTAC 2541
 Db 2568 CATATTAAGAGCAAAATTAATCTGAGAGAAATTTTAAAAATGTTTTGTTCTTAC 2627
 QY 2542 AAATGAGATTTTGTGATGAGCTTACCACTTTTCAAGCAATTTATTAATGAGAAAT 2601
 Db 2628 AAATGAGATTTTGTGATGAGCTTACCACTTTTCAAGCAATTTATTAATGAGAAAT 2687
 QY 2602 TAACTTACTCAAGCAATGATTTGAAGGAAAGTGCATATTAATGAGATGATTAATGTT 2661
 Db 2688 TAACTTACTCAAGCAATGATTTGAAGGAAAGTGCATATTAATGAGATGATTAATGTT 2747
 QY 2662 GTGTGCGAGCTGCTGCCCAATCAATTTCTTAACATGAGCTCCAGTTACCTAATGTT 2721
 Db 2748 GTGTGCGAGCTGCTGCCCAATCAATTTCTTAACATGAGCTCCAGTTACCTAATGTT 2806
 QY 2722 TCACTGACCAAGAGATGATTTACCTCACTGATGATGATGATGATGATGATGATGATG 2781
 Db 2807 TCACTGACCAAGAGATGATTTACCTCACTGATGATGATGATGATGATGATGATGATG 2866
 QY 2782 ACTGACATGAGATTAATGATGCTGATGATGATGATGATGATGATGATGATGATGATG 2841
 Db 2867 ACTGACATGAGATTAATGATGCTGATGATGATGATGATGATGATGATGATGATGATG 2925
 QY 2842 ACAATTTGCCATTAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2901
 Db 2926 ACAATTTGCCATTAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2985
 QY 2902 AAAGTCAAGACCTTGGCTCAAGAGAGCTGTGAGCCAGATTCATTTAAGAGAAAT 2958
 Db 2986 AAATCTC-----TGTCTCAAGAGAGCTGTGAGCCAGATTCATTTAAGAGAAAT 3034

RESULT 2
 AY412347
 LOCUS
 DEFINITION
 Homo sapiens ESRG gene, VIRUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION
 AY412347
 VERSION
 AY412347.1 GI:39768312
 KEYWORDS
 GSS.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1377)
 Clark,A.G., Glawoski,S., Nielson,D., Thomas,P., Kejarival,A.,
 Todd,M.A., Tenenbaum,D.M., Cavellio,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,

TITLE Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1377)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1377
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1377
/gene="ESRRG"
/locus_tag="HCM4527"

Query Match 46.1%; Score 1375.4; DB 29; Length 1377;
Best Local Similarity 99.9%; Pred. No. 9.3e-275;
Matches 1376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

gene

126 ATGATTTGGTGAAGCTTTGCTTCTGATTTCTTTTCCCTGACATGAGGAAGAGCTT 185
1 ATGGATTTGGTGAAGCTTTGCTTCTGATTTCTTTTCCCTGACATGAGGAAGAGCTT 60
186 CTCTGCAAGATGTCAACAAAGATGACACATTGATTCAGCTTTCCTTCAATGAC 245
61 CTCTGCAAGATGTCAACAAAGATGACACATTGATTCAGCTTTCCTTCAATGAC 120
246 ACGGAACCTTCCAGCCAGCTTCCCTGACGGAACAGCTCAACACAGCCCTGTGAC 305
121 ACGGAACCTTCCAGCCAGCTTCCCTGACGGAACAGCTCAACACAGCCCTGTGAC 180
306 TCTTCAGAGCCAGAGGAGAGCTGACGTTCAACATGATGGCCATCAAGAGAGCTTAC 365
181 TCTTCAGAGCCAGAGGAGAGCTGACGTTCAACATGATGGCCATCAAGAGAGCTTAC 240
366 TCGCCACTCTCTACCTTCTGCTCTATCTCTGAGGAGGATGAGGAGCTTCAAGAACTG 425
241 TCGCCACTCTCTACCTTCTGCTCTATCTCTGAGGAGGATGAGGAGCTTCAAGAACTG 300
426 TATGATGACTGCTCCAGCAGCAATTTGTAAGATCCCAAGCCAAAGTGAATACATGCTC 485
301 TATGATGACTGCTCCAGCAGCAGCAATTTGTAAGATCCCAAGCCAAAGTGAATACATGCTC 360
486 AACTCGATGCCAGAGAGCTGTTTGAATGTGTGTGTGATCATCGCTTCTGGGTACCATAT 545
361 AACTCGATGCCAGAGAGCTGTTTGAATGTGTGTGTGATCATCGCTTCTGGGTACCATAT 420
546 GGGGTAGCATCATGTGAAGCTTGAAGGATTTCTTCAAGAGAGCAATTCAGAGCAATTA 605
421 GGGGTAGCATCATGTGAAGCTTGAAGGATTTCTTCAAGAGAGCAATTCAGAGCAATTA 480
606 GAATACAGCTGCTTCCGCAAGATGATGTGAATTCACAAAGCCGACGCTAAATCTCTGC 665
481 GAATACAGCTGCTTCCGCAAGATGATGTGAATTCACAAAGCCGACGCTAAATCTCTGC 540
666 CAGGCTTGGCTTCAATGAAGTGTAAAGTGGGATGCTGAAGAGAGGGGTGCTTT 725
541 CAGGCTTGGCTTCAATGAAGTGTAAAGTGGGATGCTGAAGAGAGGGGTGCTTT 600
726 GACAGAGTAGTGAGGTGCGCAGAAATGACAAAGGAGAGATGATGGGAGAACAGGCCA 785
601 GACAGAGTAGTGAGGTGCGCAGAAATGACAAAGGAGAGATGATGGGAGAACAGGCCA 660

QY 786 TACCTGAACCTCAGCTGGTTGAGCCAGCCAAAGACCATATACAGATTTCTCAAT 845
DB 661 TACCTGAACCTCAGCTGGTTGAGCCAGCCAAAGACCATATACAGATTTCTCAAT 720
QY 846 TTGTTGGTGGTGAACCCGAGAAAGATCTATGCGATGCTGACCTCTGCTCCCGACAT 905
DB 721 TTGTTGGTGGTGAACCCGAGAAAGATCTATGCGATGCTGACCTCTGCTCCCGACAT 780
QY 906 GACATCAAGCCCTCACTACCTGTGTGACTTGGCCGACGAGAGTTGGTGTATCAT 965
DB 781 GACATCAAGCCCTCACTACCTGTGTGACTTGGCCGACGAGAGTTGGTGTATCAT 840
QY 966 GGATGGGCGAGCAATATTCAGCTTCCCTGAGGCTGCTTCCCTGAGGACCATGACCTT 1025
DB 841 GGATGGGCGAGCAATATTCAGCTTCCCTGAGGCTGCTTCCCTGAGGACCATGACCTT 900
QY 1026 CTGCAAGAGTCTTGAATGAGAAATTTGATCTTGGTGTGTATACCGGTCTCTTCAATT 1085
DB 901 CTGCAAGAGTCTTGAATGAGAAATTTGATCTTGGTGTGTATACCGGTCTCTTGGTTT 960
QY 1086 GAGGATGAAGCTTGTCTATGACAGAGATTTATATATGAGCAAGACCATGCAATATGCA 1145
DB 961 GAGGATGAAGCTTGTCTATGACAGAGATTTATATATGAGCAAGACCATGCAATATGCA 1020
QY 1146 GGCTTCTTGAATCTAATATATGCTATCTGACGCTGTGTAAGAAATGCAAGAGCATGAG 1205
DB 1021 GGCTTCTTGAATCTAATATATGCTATCTGACGCTGTGTAAGAAATGCAAGAGCATGAG 1080
QY 1206 CTGGAAGAAAGAAATTTGTGACCTCAAGCTATAGCTTGTGATTAATTCAGATCCATG 1265
DB 1081 CTGGAAGAAAGAAATTTGTGACCTCAAGCTATAGCTTGTGATTAATTCAGATCCATG 1140
QY 1266 CACATATGAAGATGTGAAGCCGTTGAGAAGCTTCAAGATGTCTTAATCATGAGAGCGTGCAG 1325
DB 1141 CACATATGAAGATGTGAAGCCGTTGAGAAGCTTCAAGATGTCTTAATCATGAGAGCGTGCAG 1200
QY 1326 GATTATGAAGCTGGGCGACACATGGAAGACCTGTGAGGCTGGCAAGATGTGATGCA 1385
DB 1201 GATTATGAAGCTGGGCGACACATGGAAGACCTGTGAGGCTGGCAAGATGTGATGCA 1260
QY 1386 CTGCCACTCTGAGGAGAGCTTACCAAGGCGCTGACAGATTTCTAACAATCAAACTA 1445
DB 1261 CTGCCACTCTGAGGAGAGCTTACCAAGGCGCTGACAGATTTCTAACAATCAAACTA 1320
QY 1446 GAAGGCAAGTCCCAATGACAACTTTTGGAAATGTGGAGGCCAAGGCTTGA 1502
DB 1321 GAAGGCAAGTCCCAATGACAACTTTTGGAAATGTGGAGGCCAAGGCTTGA 1377

RESULT 3
AY412348 1377 bp DNA linear GSS 16-DEC-2003
LOCUS
DEFINITION Pan troglodytes ESRRG gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
VERSION
AY412348.1 GI:39768313
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1377)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES
 source
 Location/Qualifiers
 1..1377

gene
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>1377
 /gene="ESRAG"
 /locus_tag="HCW4527"

ORIGIN

Query Match 45.2%; Score 1349.6; DB 29; Length 1377;
 Best Local Similarity 98.4%; Pred. No. 2.1e-269;
 Matches 1355; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 126 ATGAGTGGTGAAGCTTTCCTCTGTAATCTTTTCCCTGCACATACAGAGAGAGCTT 185
 Db 1 ATGGATTCGGTGAAGCTTTCCTCTGTAATCTTTTCCCTGCACATACAGAGAGAGCTT 60
 QY 186 CTTCGAGAGATGTCAAAAGAGATGACACATTTGATTCAGCTGTTCCTTCATCAG 245
 Db 61 CTCTCAGAGATGTCAAAAGAGATGACACATTTGATTCAGCTGTTCCTTCATCAG 120
 QY 246 AGGAAACCTTCAGGCCACCTCCCTGAGGAGACAGCTCAACACACAGCCCTGGTGGC 305
 Db 121 AGGAAACCTTCAGGCCACCTCCCTGAGGAGACAGCTCAACACACAGCCCTGGTGGT 180
 QY 306 TCTTGAAGCGCAGATGGAGACTACAGTTCAACCATGATGGCCATCAGACGAGACTTGA 365
 Db 181 TGTGACAGAGCCAGTGGAGCTACAGTTCAACCATGATGGCCATCAGACGAGACTTGA 240
 QY 366 TGGCACTCTTCAACCTTCTGCTCTATCTGAGAGGTAGTGGGCTTCAAGAACTG 425
 Db 241 TCGNNNNNTCTCAACCTTCTGCTCTATCTGAGAGGTAGTGGGCTTCAAGAACTG 300
 QY 426 TATGATGATGCTGCCAGACCATTTGTGAAGATCCCAAGTGTGAATATCATGCTC 485
 Db 301 TATGATGATGCTGCCAGACCATTTGTGAAGATCCCAAGTGTGAATATCATGCTC 360
 QY 486 AACTGATGCCCAAGAGACTGTGTGTAGTGTGTGATGATGCTTCTGGTACCATAT 545
 Db 361 AACTGATGCCCAAGAGACTGTGTGTAGTGTGTGATGATGCTTCTGGTACCATAT 420
 QY 546 GGGGTAGCATCATGTGAAGCTTGCAAGGCACTTTCTTAAGAGCAATTCAGGCAATATA 605
 Db 421 GGGGTAGCATCATGTGAAGCTTGCAAGGCACTTTCTTAAGAGCAATTCAGGCAATATA 480
 QY 606 GAATAGAGTCCCTCCAGACCATGTAATGTGAATCAAAAGGCGAGAGCTTAATCTCG 665
 Db 481 GAATAGAGTCCCTCCAGACCATGTAATGTGAATCAAAAGGCGAGAGCTTAATCTCG 540
 QY 666 CAGGCTTGCCTTCATGAAGTGTATAAAGTGGGATGCTGAAAAGGGGCTGCTCTT 725
 Db 541 CAGGCTTGCCTTCATGAAGTGTATAAAGTGGGATGCTGAAAAGGGGCTGCTCTT 600
 QY 726 GACAGAGTACGTGAGGTGGCGAAGATACAGCGCAGATAGATGGGAGAACAGCCCA 785
 Db 601 GACAGAGTACGTGAGGTGGCGAAGATACAGCGCAGATAGATGGGAGAACAGCCCA 660
 QY 786 TACCTGAACCTTCAGCTGTTCAGCGACCAAAAAGCATATATACAGATGTGCTCAT 845
 Db 661 TACCTGAACCTTCAGCTGTTCAGCGACCAAAAAGCATATATACAGATGTGCTCAT 720
 QY 846 TTGTTGGTGGCTGAACCGAGAGATCTATGCTGACCTGACCTACTGTCCTCCGACAGT 905
 Db 721 TTGTTGGTGGCTGAACCGAGAGATCTATGCTGACCTGACCTACTGTCCTCCGACAGT 780

QY 906 GACATCAAGCCCTCAGTACAGCTGTGTGAGCTTGGCCGACCGAGAGTGGTGTATCAT 965
 Db 781 GACATCAAGCCCTCAGTACAGCTGTGTGAGCTTGGCCGACCGAGAGTGGTGTATCAT 840
 QY 966 GGAATGGGGAAGCATATTCAGAGCTTTCACAGCTGTCTCGGCGAGACAGTGAAGCTT 1025
 Db 841 GGAATGGGGAAGCATATTCAGAGCTTTCACAGCTGTCTCGGCGAGACAGTGAAGCTT 900
 QY 1026 CTGCAGATGCTTGGATGGAATTTTGTATCCTGTGTGTGTATACCGGTCTCTTCATT 1085
 Db 901 CTGCAGATGCTTGGATGGAATTTTGTATCCTGTGTGTGTATACCGGTCTCTTCATT 960
 QY 1086 GAGATGAACCTTGTCTATGACAGACATATATATATGACAGAGACAGACCAATATGCA 1145
 Db 961 GAGATGAACCTTGTCTATGACAGACATATATATATGACAGAGACAGACCAATATGCA 1020
 QY 1146 GGCCTTCTTGAATCTAATATATGCTATCTGACAGCTGTATAGAAATATACAGAGATGA 1205
 Db 1021 GGCCTTCTTGAATCTAATATATGCTATCTGACAGCTGTATAGAAATATACAGAGATGA 1080
 QY 1206 CTGAAAAAAGAAATTTGTCAACCTCAAGCTATAGCTCTTGTATTCAGATTCATG 1265
 Db 1081 CTGAAAAAAGAAATTTGTCAACCTCAAGCTATAGCTCTTGTATTCAGATTCATG 1140
 QY 1266 CACATAGAAAGATTTGAAGCCCTTCAAGAGCTTCAGAGATGCTTACATAGAGGCTGCG 1325
 Db 1141 CACATAGAAAGATTTGAAGCTTTCAGAGCTTCAGAGATGCTTACATAGAGGCTGCG 1200
 QY 1326 GATTATGAAGCTGGCAGACATAGAAAGCCCTGCTGAGCTGCGAGATGCTGATGACA 1385
 Db 1201 GATTATGAAGCTGGCAGACATAGAAAGCCCTGCTGAGCTGCGAGATGCTGATGACA 1260
 QY 1386 CTGCACTTCGTAGGAGACACTTACCAAGCCGTGACAGATTTCTACATCAAACTA 1445
 Db 1261 CTGCACTTCGTAGGAGACACTTACCAAGCCGTGACAGATTTCTACATCAAACTA 1320
 QY 1446 GAAGGGAAGTCCCAATGACAAACTTTTGTGAAGTGTGGAGCCAGGCTGA 1502
 Db 1321 GAAGGGAAGTCCCAATGACAAACTTTTGTGAAGTGTGGAGCCAGGCTGA 1377

RESULT 4
 AY412349 1377 bp DNA linear GSS 16-DEC-2003
 LOCUS
 DEFINITION
 Mus musculus ESRAG gene, VIRUAL TRANSCRIPT, partial sequence.
 ACCESSION
 AY412349
 VERSION
 AY412349.1 GI:39768314
 KEYWORDS
 GSS.

SOURCE
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
 Adams, M.D. and Cargill, M.
 1 (bases 1 to 1377)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios

JOURNAL
 PUBLISHED
 14671302
 2 (bases 1 to 1377)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 Location/Qualifiers

COMMENT
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 Location/Qualifiers

source 1..1377
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>1377
 /gene="ESRNG"
 /locus_tag="HGM4527"

Query Match 41.2% Score 1229.8; DB 29; Length 1377;
 Best Local Similarity 93.3%; Pred. No. 1.5e-244;
 Matches 1285; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 126 ATGATTTGGTAGAACTTTGCTTCTGATCTTTTCCCTGCACTACGAGAGAGCTT 185
 Db 1 ATGATTTGGTAGAACTTTGCTTCTGATCTTTTCCCTGCACTACGAGAGAGCTT 60
 QY 186 CTCTGCAAGATGTCAAACAAAGATGACACATTTGCTTCCAGCTGTTCTCTCATCAAG 245
 Db 61 CTCTGCAAGATGTCAAACAAAGATGACACATTTGCTTCCAGCTGTTCTCTCATCAAG 120
 QY 246 ACGGAACCTTCCAGCCGACCTCCCTGACGAGACAGCTCAACCAAGCCCTGCTGGC 305
 Db 121 ACGGAACCTTCCAGCCGACCTCCCTGACGAGACAGCTCAACCAAGCCCTGCTGGC 180
 QY 306 TCTTTGAGACGCGAGTGGAGCTACAGTTCAACATGAAATGGCCATCAGAACGACTTAC 365
 Db 181 TCTTTGAGACGCGAGTGGAGCTACAGTTCAACATGAAATGGCCATCAGAACGACTTAC 240
 QY 366 TCGCCACTCTCTACCTCTTCTGCTCTTACTCTGGAGAGTATGGGCTCTGACAGAACTG 425
 Db 241 TCGCCACTCTCTACCTCTTCTGCTCTTACTCTGGAGAGTATGGGCTCTGCTCGGAACTG 300
 QY 426 TATGATGACTGCTCCAGACCACTATTTGAAGATCCCGACCAAGTGAATACATGCTC 485
 Db 301 TATGATGACTGCTCCAGACCACTATTTGAAGATCCCGACCAAGTGAATACATGCTC 360
 QY 486 AACTGATGCCCAAGAGACTGTGTTAGTGTGTGAGTGCATGCTGTTGGGTACCACTAT 545
 Db 361 AACTGATGCCCAAGAGACTGTGTTAGTGTGTGAGTGCATGCTGTTGGGTACCACTAT 420
 QY 546 GGGGTAGATGATGTGAAGCTGCAAGGCACTTTTCAAGAGAGCAATTCAGGCAATTA 605
 Db 421 GGGGTAGATGATGTGAAGCTGCAAGGCACTTTTCAAGAGAGCAATTCAGGCAATTA 480
 QY 606 GAATACAGCTGCTCCGACCAAGATGAATGTGAATCAACAAAGCGACGTAATTCCTGC 665
 Db 481 GAATACAGCTGCTCCGACCAAGATGAATGTGAATCAACAAAGCGACGTAATTCCTGC 540
 QY 666 CAGGCTTCCGCTTCATGAAGTGTTTAAAGTGGGCAATGCTGAAAGAGGGGTGCTCTT 725
 Db 541 CAGGCTTCCGCTTCATGAAGTGTCTCAAGTGGGCAATGCTGAAAGAGGGGTGCTCTT 600
 QY 726 GACGAGTACGTGAGAGTGGCGCAAGATGACAGCGAGATGATGCGGAGAAACAGCCCA 785
 Db 601 GACGAGTACGTGAGAGTGGCGCAAGATGACAGCGAGATGATGCGGAGAAACAGCCCA 660
 QY 786 TACCTGAACCTCAGCTGTGTTCAAGCGCCAAAAGCCATATACAGATTTGTCTCAT 845
 Db 661 TACCTGAACCTCAGCTGTGTTCAAGCGCCAAAAGCCATATACAGATTTGTCTCAT 720
 QY 846 TTGTTGATGCTGAACCGAGAGAGATCTATGCACTGACCTTACTGTCCTCCAGAGT 905
 Db 721 TTGTTGATGCTGAACCGAGAGAGATCTATGCACTGACCTTACTGTCCTCCAGAGT 780
 QY 906 GACATCAAGGCTTCATACCTGTGATGCTTGGCCACCGAGAGTGTGTATCATTT 965
 Db 781 GACATCAAGGCTTCATACCTGTGATGCTTGGCCACCGAGAGTGTGTATCATTT 840
 QY 966 GATGGGCGAAGCATATTCAGAGCTTCTCAACGCTGTCTCTGGGCGACAGATGAGCTT 1025
 Db 841 GATGGGCGAAGCATATTCAGAGCTTCTCAACGCTGTCTCTGGGCGACAGATGAGCTT 900

QY 1026 CTGACAGTGTGCTGATGAGAAATTTTGAATCCTTGATGTGTCGATACCGGCTCTTCAATT 1085
 Db 901 CTGACAGTGTGATGATGAGAAATTTTGAATCCTTGATGTGTCGATACCGGCTCTTCAATT 960
 QY 1086 GAGATGAACCTTGTCTATGACAGACGATTAATATATGACGAGAACCAATTCAAATTAGCA 1145
 Db 961 GAGATGAACCTTGTCTATGACAGACGATTAATATATGATGAGAACCAATTCAAATTAGCA 1020
 QY 1146 GGCCTTCTGATCTTAAATATATGCTATCTGACGCTGTGTAAGAAATACAAAGCATGAG 1205
 Db 1021 GGCCTTCTGATCTTAAATATATGCTATCTGACGCTGTGTAAGAAATACAAAGCATGAG 1080
 QY 1206 CTGAAAAAAGAAATTTGTACCTCAAGCTATAGCTCTTGCTATTCAGACTCCATG 1265
 Db 1081 CTAGAGAAAGAAATTTGTATGCTACCTCAAGCATAGCTCTTGCTATTCAGACTCCATG 1140
 QY 1266 CACATGAAAGTGTGAAAGCCGTTGAGAGCTTGAGATGCTTATCATGAAAGCGCTGAC 1325
 Db 1141 CATATGAAAGTGTGAAAGCTGTGACAGAACTTGAGATGCTTATCATGAGGCTGAC 1200
 QY 1326 GATTATGAGTGTGCGCCAGACATAGAAAGCCCTGCTGAGCTGGCAATGCTATGACA 1385
 Db 1201 GATTATGAGTGTGCGCCAGACATAGAAAGCCCTGCTGAGCTGGCAATGCTATGAC 1260
 QY 1386 CTGCCACTCTGAGAGCAGACCTCTACCAAGCCGCTGACGATTTCTACATCAACTAACTA 1445
 Db 1261 CTGCCACTCTGAGAGCAGACCTCTACCAAGCCGCTGACGATTTCTACATCAACTAACTC 1320
 QY 1446 GAAGGCAAAAGTCCCAATCACAATCTTTTGGAAATGTTGAGAGCCAGAGCTCTGA 1502
 Db 1321 GAAGGCAAAAGTCCCAATCACAATCTTTTGGAAATGTTGAGAGCCAGAGCTCTGA 1377

RESULT 5
 AL700159
 LOCUS 893 bp mRNA linear EST 04-SEP-2003
 DEFINITION DKFZp686J14117.1 893 bp (synonym: h1cc3) Homo sapiens CDNA clone
 ACCESSION DKFZp686J14117.5', mRNA sequence.
 VERSION AL700159
 KEYWORDS AL700159.1 GI:19620692
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 893)
 Ottewaelder,B., Obermaier,B., Mewes,W., Mewes,H.W., Weil,B. and
 Wiemann,S.
 TITLE EST (Ottewaelder,B., Obermaier,B., Mewes,H.W., Weil,B. and
 Wiemann,S.)
 JOURNAL Unpublished (2001)
 COMMENT Contact: MIPS

MTBS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Medigenomix (Martinsried/Germany) within the cDNA
 sequencing consortium of the German Genome Project. No si sequence
 available.
 This clone (DKFZp686J14117) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 Location/Qualifiers

FEATURES
 source
 1..893
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686J14117"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_id="686 (synonym: h1cc3)"
 /note="Vector: pTriplEx2; Site_1: SfIA; Site_2: SfIB;

ORIGIN

cDNA-collection"

Query Match 27.6%; Score 823.4; DB 9; Length 893;
 Best Local Similarity 98.7%; Pred. No. 3e-160;
 Matches 883; Conservative 0; Mismatches 6; Indels 6; Gaps 5;

```

QY 1447 AAGCAAAAGTCCCAATGACAAATCTTTTGGAAATGTTGAGAGCCCAAGTCTACTTA 1506
DB 1 AAGGCAAAAGTCCCAATGACAAATCTTTTGGAAATGTTGAGAGCCCAAGTCTACTTA 60
QY 1507 AAGTCCCTGGGCTTCCCATCTCTCATGTTTAAAGGAAATTAACCCCAAGGTAT 1566
DB 61 AAGTCCCTGGGCTTCCCATCTCTCATGTTTAAAGGAAATTAACCCCAAGGTAT 120
QY 1567 GTGCAAGAACTTAGAGTTAGTTAAACAATCAAAAATCAACAGACTGCACTGAAT 1626
DB 121 GTGCAAGAACTTAGAGTTAGTTAAACAATCAAAAATCAACAGACTGCACTGAAT 180
QY 1627 TAGGCAAGAACTTAGAGTTAGTTAAACAATCAAAAATCAACAGACTGCACTGAAT 1685
DB 181 TAGGCAAGAACTTAGAGTTAGTTAAACAATCAAAAATCAACAGACTGCACTGAAT 240
QY 1686 CTAATTTCTCCCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1745
DB 241 CTAATTTCTCCCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 1746 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1805
DB 301 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 1806 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1865
DB 361 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
QY 1866 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1925
DB 421 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 478
QY 1926 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1985
DB 479 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 538
QY 1986 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2045
DB 539 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 598
QY 2046 AAAAGAGGTCTTAAGTGAAGAGATGAAAGCAGCCCTGCAAGAGAT- GGAAGTCCA 2104
DB 599 AAAAGAGGTCTTAAGTGAAGAGATGAAAGCAGCCCTGCAAGAGATGGAAGTCCA 658
QY 2105 TAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2164
DB 659 TAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 718
QY 2165 AAGAGAGAGAACCAAGCTTCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2223
DB 719 AAGAGAGAGAACCAAGCTTCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 778
QY 2224 GTATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2282
DB 779 GTATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 838
QY 2283 GCTTTCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2337
DB 839 GCTTTCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 893

```

RESULT 6
 BX369768 919 bp mRNA EST 08-MAY-2003
 LOCUS BX369768 Homo sapiens NEUROBLASTOMA COR 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DC014YD14 5-PRIME, mRNA sequence.

ACCESSION BX369768
 VERSION BX369768.1 GI:30441500
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 919)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 904.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BA007ZG02_CS00591_1&cluster=904.r.
 Contact : Feng Liang Email : fliang@life.com URL :
 http://fulllength.invitrogen.com/invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BA007ZG02_CS00591_1.

FEATURES

source

1..919
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC014YD14"
 /tissue_type="NEUROBLASTOMA COR 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 25.6%; Score 763; DB 13; Length 919;
 Best Local Similarity 98.6%; Pred. No. 1e-147;
 Matches 799; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

```

QY 2145 CCCAATGACTAAGGATCAAGAGAGAGAACCAAGCTTCTTAAGATGAGTCAACATA 2204
DB 36 CCCAAGACTAAGGATCAAGAGAGAGAACCAAGCTTCT- AAAGTACAGTCAACATA 94
QY 2205 TACAATTGACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2264
DB 95 TACAATTGACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 154
QY 2265 AGCAAGTTCATCGGCTGCTTCTTATCATGCTTTTCCATCTAAGTCAAGTCAAGCAT 2324
DB 155 AGCAAGTTCATCGGCTGCTTCTTATCATGCTTTTCCATCTAAGTCAAGTCAAGCAT 214
QY 2325 TTGATTCCTTAATGTTTTTTCAGGCTCCAGGATTTGTTAGTTAGTACTATGTA 2384
DB 215 TTGATTCCTTAATGTTTTTTCAGGCTCCAGGATTTGTTAGTTAGTACTATGTA 274
QY 2385 CTTTTTCAGGGAATAGTTTAAGCTTATTCATTCATCACTAAGAGAAATTAAGAT 2444
DB 275 CTTTTTCAGGGAATAGTTTAAGCTTATTCATTCATCACTAAGAGAAATTAAGAT 334
QY 2445 ACTGCAATTTTGCTGCTTGAACAATTAACAACAATTAACAACAATTAACAACAATTAACA 2504
DB 335 ACTGCAATTTTGCTGCTTGAACAATTAACAACAATTAACAACAATTAACAACAATTAACA 394
QY 2505 GAGGAAGATTTTAAAGATTTTGTCTTCTCTCAATTAAGAGATTTTGTGACAG 2564
DB 395 GAGGAAGATTTTAAAGATTTTGTCTTCTCTCAATTAAGAGATTTTGTGACAG 454
QY 2565 CTTTACCCTTTTCAGCCATTTATTAATATGGAATTTAACTTACTCAAGCAATAGTTGA 2624
DB 455 CTTTACCCTTTTCAGCCATTTATTAATATGGAATTTAACTTACTCAAGCAATAGTTGA 514

```

QY	2635	AGGGAAGTGCATATTTATCACGAGATGCAATTTATGTGTGTGGCAAGCTGGTCCCAAC	2683
Dp	515	AGGGAAGTGCATATTTATCACGAGATGCAATTTATGTGTGTGGCAAGCTGGTCCCAAC	574
QY	2685	TCATTTTCTTAACTAGAGCTCCAGTTTACCTPAATGTTCAGTGCACAAAGATGAGATT	2744
Dp	575	TCATTTTCTTAACTAGAGCTCCAGTTTACCTPAATGTTCAGTGCACAAAGATGAGATT	634
QY	2745	ACACCTACAGTACCTCTGAGTAGTACATATATATAGCACTGCACATGAGATATAGATCCG	2804
Dp	635	ACACCTACAGTACCTCTGAGTAGTACATATATATAGCACTGCACATGAGATATAGATCCG	694
QY	2805	TAGAATTGTCAAGAGTGCACCTCTCTACTTTGGAGGTACAAATGGCATATGATTTCTAGC	2864
Dp	695	TAGAATTGTCAAGAGTGCACCTCTCTACTTTGGAGGTACAAATGGCATATGATTTCTAGC	754
QY	2865	TGCCATGGTGTAGGATGTGATACACTGTTTGGCAAAATGCACAGACTTGGCTCAGA	2924
Dp	755	TGCCATGGTGTAGGATGTGATACACTGTTTGGCAAAATGCACAGACTTGGCTCAGA	812
QY	2925	AGGAGCTGTGGGCCAGTATTCATTTAAG	2954
Dp	813	AGGAGCTGTGGGCCAGTATTCATTTAAG	842

RESULT 7	
EX332672	
LOCUS	BX332672
DEFINITION	1005 bp mRNA linear EST 01-MAY-2003
	BX332672 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
	cDNA clone CS0DC014YD14 5'-PRIME, mRNA sequence.
ACCESSION	U733267.3

ACCESSION	BM332672	
VERSION	BM332672.1	GI:30312142
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryote; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1. (bases 1 to 1005)	
AUTHORS	L1, W. B., Gruber, C., Jessee, J., and Polayes, D.	
TITLE	Full-length cDNA libraries and normalization	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Genoscope	

Genoscope - Centre National de Séquençage
BP 191 91066 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 904.r For
more information about this cluster, see
<http://www.genoscope.cns.fr/>
cgl-bam/cluster.cgi?seq=CSODC014DB07QP1&cluster=904.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>
Paradey Avenue Genoscope sequence ID : CSODC014DB07QP1.
Location/Qualifiers

```

source
1.1005
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC014YD14"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-0190(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

	ORIGIN
Query Match	24.9%; Score 744.6; DB 13; Length 1005;
Best Local Similarity	88.6%; Pred. No. 6.6e-144;
Matches 820; Conservative 43; Mismatches 53; Indels 9; Gaps 6;	

QY 1640 ATGAAGCAGCTTTGAGATTCCTCCATAGGTTCTGTAGAG-TTCTTTCTACTTTCTCCAT 1698
|||||

Db	64	ATGAAGACGCTTTAGATTCCTCCATAGAGTTCCTGATGAGTTCTTTCCTACTTTCCTCAT	123
Qy	1699	CATCTCTCTTCTCTTCTCTTCCCAATTCCTTCTCTCTATTTTCTCTCTCTCTCTC	1758
Db	124	CATCTCTCTCTCTCTCTCTTCCCAATTCCTCTCTCTCTATTTTCTCTCTCTCTC	183
Qy	1759	TTTCACTCTCTTATTTCTTCTGCTCTTCTTCAATTCCTAGTTCCTCATTTCTTATTTCT	1818
Db	184	TTTCACTCTCTTATTTCTTCTGCTCTTCTTCAATTCCTAGTTCCTCATTTCTCTTATTTCT	243
Qy	1819	TCCGGTCGCGCGCTCTCTCTTCTCTTAACTGACCTCAATTCCTGCTCTTCTCAACG	1878
Db	244	TCCGGTCGCGCTCTCTCTTCTTCTTAACTGACCTCAATTCCTGCTCTTCTCAACG	303
Qy	1879	TTCCCTCTTTTCTTAAATTGAAATAGCTTAGTTAAAAAATAAATCTCCCTCTGCC	1938
Db	304	TTCCCTCTTTTCTTAAATTGAAATAGCTTAGTT--AAAAAATAATCCCTCTTCCG	361
Qy	1939	CTTTCCTCTTCCCTTCTTCTCTCTTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTT	1998
Db	362	CTTTCCTCTTCCCTTCTTCTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTT	421
Qy	1999	GACCTCTCTTCCATCTTCTTCTTCTTCTCTGCTGCTGTAACCTTTTAAAGAGTCTCT	2058
Db	422	GACCTCTTCTTCCATCTTCTTCTTCTTCTCTGCTGCTGTAACCTTTTAAAGAGTCTCT	481
Qy	2059	AACCTGAAAGAGATGAGAACCCAGCTCTGCCAAGAGATGAGATCATATATATGATGCCA	2118
Db	482	AACCTGAAAGAGATGAGAACCCAGCTCTGCCAAGAGATGAGATCATATATATGATGCCA	541
Qy	2119	GTGAAC-TTATATGTGAACCATACCGTCTCCCAATGACTTAAGGATTCAAAGAGAACCA	2177
Db	542	GTGAACATTTATGTGAACCAACCGTCTCCCAATGACTTAAGGATTCAAAGAGAACCA	601
Qy	2178	ACGTTCTCTAAAGATGAGAGTGCACATATCAAAATGACTGAGTGCATATATGATTCAT	2237
Db	602	ACGTTCTCTAAAGATGAGAGTGCACATATCAAAATGACTGAGTGCATATATGATTCAT	661
Qy	2238	GGAGGACGCTCTTAATATGACAACCTTAAGCAACCTTGCATGCGCTCTTATCATCTGC	2297
Db	662	GGAGGACGCTCTTAATATGACAACCTTGCATGCGCTCTTATCATCTGC	721
Qy	2298	TTTTTCATCTAGATCACTTACAGCCATTTGATTCCTAATTTGTTTTTCAAGTCTCCAG	2357
Db	722	TTTTTCATCTAGATCACTTACAGCCATTTGATTCCTAATTTGTTTTTCAAGTCTCCAG	781
Qy	2358	GTATTTGTAGTTAGTACTATGTATCTTTTCAGGAAATAGTTTAAGCTTATTCATTT	2417
Db	782	GTATTTGTAGTTAGTACTATGTATCTTTTCAGGAAATAGTTTAAGCTTATTCATTT	841
Qy	2418	CATGCAATACTAAAGAGAAATAGAAATACGTCAATTTTGTCTGCTGCTTGAACAATTAAG	2477
Db	842	TAKCSAA-ANPAAAGMMMAAAAAAANAAGCAA-WTTGGCGCTKGCTTTCMAAAATTAAG	899
Qy	2478	AACATATATGAGGACAAATGAATCCTGAAAGGAAGATTTTAAATAATGTTTTGTTCTTC	2537
Db	900	CAAAATATATCAGGCAAAAAGCAATTCWCAASCAHATTTT--AAAKYTTTXYTTTTT	956
Qy	2538	TTTCAAAATGAGATTTTTTTTGTACC	2562
Db	957	TTTAAAKGCGMMTTTTTXXTTAAAC	981

RESULT	8
LOCUS	AY409734
DEFINITION	AY409734 1296 bp DNA linear GSS 16-DEC-2003
ACCESSION	Homo sapiens ESRB gene, VIRUTAL TRANSCRIPT, partial sequence,
VERSION	AY409734
KEYWORDS	Genomic survey sequence.
SOURCE	AY409734.1 GI:39765702
ORGANISM	GSS.
	Homo sapiens (human)
	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
ORIGIN
Query Match
Best Local Similarity
Matches
23.6%; Score 704; DB 29; Length 1296;
71.9%; Pred. No. 1.7e-135;
936; Conservative 0; Mismatches 360; Indels 6; Gaps 1;
195
1
255
61
315
121
375
180
435
235
495
295
555
355
615
415
675
475
735
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1296)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Interfing nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1296)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..1296
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1296
/gene="ESRRB"
/locus_tag="HCM3656"

Db 535 CGTGGAGGCCGTCAGAAATACAGAGAGCGGCTGACCTCAGAGAGGCCCATACCTGAGC 594
Qy 795 CCGTAGCTGTTCAAGCCAGCCAAAGCCATTAACAAGATTGTCATTTGGTG 854
Db 595 TTACAAATTTCTCCACCTGCTAAAGAGCCATTGACCAAGATTGCTCATCTACTCTGG 654
Qy 855 GCTGAACCGGAGAAATCTATAGCCATGCTGACCTTACCTGCCCCGAGAGTACATCAA 914
Db 655 GCTGACCGGACCAAGCTTATGCAATGCTGCTCCCTGATGCTGAGGGGACATCAAG 714
Qy 915 GCCCTCATCACTGTGACTTGGCCAGCAGAGAGTGTGTATCATTTGATGATGAGTGG 974
Db 715 GCCCTGACCATCTCTGTGACTTGGCCAGCAGAGAGTGTGTATCATTTGCTGAGTGG 774
Qy 975 AAGCATATTCAGAGCTTCTCCACGCTGTGCTGCGGACACAGATAGAGCTTTGACAGT 1034
Db 775 AAGCATATTCAGAGCTTCTCCACGCTGTGCTGCGGACACAGATAGAGCTTTGACAGT 834
Qy 1035 GCTTGATGAGAAATTTTATCTGTTGTCGATACCGGCTCTTTTCATTTGATGAGATGA 1094
Db 835 GCTGATGAGAAATCTTATCTGCTGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
Qy 1095 CTTGTCTATGACAGATTAATTAATGACAGAGACAGTCCAAATTGACAGGCTTCTT 1154
Db 895 CTGATGTACGCTGAGAGACTACATCATGATGAGAGCACTCCGCTGCGGGGCTGCTG 954
Qy 1155 GATCTAATTAATCTATCTGCTGAGAGTGTGAAGAAATACAGAGATGAAGTGTGA 1214
Db 955 GAGCTCTACCGGCGCATCTGCACTGATGACAGAGTGTGAAGAAATCTGAGAGAG 1014
Qy 1215 GAAAGATTTGTACCTCCTAAGCTTATGCTTGTGCTTATTCAGATCCATGACATGAA 1274
Db 1015 GAGGATTTGTACCTCCTAAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 1074
Qy 1275 GATGTTGAAGCCGTTACAGAGCTTCAAGATGCTTCAATGAAGCGCTGACAGATTATGA 1334
Db 1075 GATCTAGAGGCTGTCCAGAGCTGACAGAGCTGCTGACAGAGCACTGACAGAGCTAG 1134
Qy 1335 GCTGCGACACATGAGAGACCTTGTGAGCTGCTGAGCTGCTGAGATGTGATGACTGCT 1394
Db 1135 CTGAGCGAGCGCATATGAGAGCCCTGAGAGAGCGGCAACCTGCTGACATGCTGCTG 1194
Qy 1395 CTGAGCGAGCGCATATGAGAGCCCTGAGAGAGCGGCAACCTGCTGACATGCTGCTG 1454
Db 1195 CTGCGGACACGCGCGCAAGCCGTGACAGCTTCTATAGCTCAAACTGACAGGCAAA 1254
Qy 1455 GTCCCATGCAAACTTTTGTGAAATGTTGAGGCGCAAG 1496
Db 1255 GTGCCATGCAAACTTTCTGAGAGATGCTGAGGCGCAAG 1296
RESULT 9
BG434768
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
799 bp mRNA linear EST 14-MAR-2001
602507976F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4604994 5',
mRNA sequence.
BG434768
BG434768.1 GI:13341274
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.femail.nih.gov
Tissue Procurement: Clontech Laboratories, Inc.
cDNA Library Preparation: Clontech Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
 Plate: ILNLT341 row: h column: 19
 High quality sequence stop: 627.
 Location/Qualifiers

FEATURES

Source

1. .799
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4604994"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_79"
 /note="Organ: placenta; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcgcc); Site 2: SfiI (ggccatctggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGACC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 23.6%; Score 703.2; DB 12; Length 799;
 Best Local Similarity 96.0%; Pred. No. 2.6e-135;
 Matches 753; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

1996 CTTGACCTCTCTTCATCTTCTTCTTCTTCTGCTGCTGAACCTTTAAAGAGCTC 2055
 1 CTTGACCTCTCTTCATCTTCTTCTTCTTCTGCTGCTGAACCTTTAAAGAGCTC 60
 2056 TCTAAGTGAAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2115
 61 TCTAAGTGAAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 2116 CCAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2175
 121 CCAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 2176 CCAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2235
 181 CCAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 2236 ATGGAG 2295
 241 ATGGAG 300
 2296 GCTTTTCATCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2355
 301 GCTTTTCATCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 2356 AGGATTTGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 2415
 361 AGGATTTGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 420
 2416 TTCAATGCAATCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2475
 421 TTCAATGCAATCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 2476 CGAACAATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2535
 481 CGAACAATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 2536 TCTTACAAATGAAG 2595
 541 TCTTACAAATGAAG 600
 2596 GGAATTTAACTTACTAGCAAGCAATGTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2654
 601 GGAATTTAACTTACTAGCAAGCAATGTTGAAGAGAGAGAGAGAGAGAGAGAGAG 660

QY 2655 TTATGTGTGTGCGCAGCTGTGCGCAACATCAATTTCTTAACATGATCCAGTTTACC 2714
 DB 661 TTATGTGTGTGCGCAGCTGTGCGCAACATTTCTTAAACATGATCCAG-TTACC 719
 QY 2715 TAAATGTTCACTGACACCAAGAGATGATTAACCTACAGTACTGATGATGATCACTA 2774
 DB 720 TAAATGTTCACTGAC-CCAAGAGTGGATTAACCTACAGGAGTCTGATGATGATTT 778
 QY 2775 TATA 2778
 DB 779 TTTA 782

RESULT 10
 LOCUS CA752448 783 bp mRNA linear EST 09-JUL-2003
 DEFINITION CA752448 UI-M-FOO-cdo-f-13-0-UI.r1 NIH_BMAP_F00 Mus musculus CDNA clone
 IMAGE:6831086 5', mRNA sequence.

ACCESSION CA752448
 VERSION CA752448.1 GI:25583085
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 783)
 NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pyx-5.

Location/Qualifiers

Source

1. .783
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6831086"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_F00"
 /note="Organ: Brain; Vector: pyx-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAAGAGGCC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemm Chin, Ph.D., program coordinator."

ORIGIN

Query Match 23.3%; Score 696.6; DB 14; Length 783;
 Best Local Similarity 93.1%; Pred. No. 6.1e-134;
 Matches 729; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY	358	GACATTGACTGGCACAACCTCTACACCTTCTGCTCCATCTCTGGAGAGATGGGCTGTCA	417
Db	1	GACTGGACTGGCACCCTCTTACCCCTCTGCTCCGATCTGGAGAGACGGGGCTGTCC	60
QY	418	GGAACCTGTATGATGACTGCTCCAGCACCATTTGTTAAAGATCCCCAGACCAAGTGTGAT	477
Db	61	GGAACCTGTATGATGACTGCTCCAGCACCATTTGTTAAAGATCCCCAGACCAAGTGTGAT	120
QY	478	ACATGCTCAACTCGATGCGCCCAAGAGACTGTGTTAAGTGTGTGTGACATGCGCTCTGGGT	537
Db	121	ATATGCTCAACTCGATGCGCCCAAGAGACTGTGTTAAGTGTGTGTGACATGCGCTCTGGGT	180
QY	538	ACCACATATGGGGTATGACATCATGTGTAAGGCTGCAAGGCACTTCTTCAAGAGACCAATTCAAG	597
Db	181	ACCACATATGGGGTATGACATCATGTGTAAGGCTGCAAGGCACTTCTTCAAGAGACCAATTCAAG	240
QY	598	GCAATATATGAATACAGCTGCGCTGCGCCAGATGAATGTGAATACACAAAGCGACAGCTGA	657
Db	241	GTAACATATGAATACAGCTGCGCTGCGCCAGATGAATGTGAATACACAAAGCGACAGCTGA	300
QY	658	AATCTCTGCAAGCTTGGCCCTTTCATGAAGTGTTTAAAGTGGGATGCTTGAAGAAAGGGG	717
Db	301	AATCTCTGCAAGCTTGGCCCTTTCATGAAGTGTTTAAAGTGGGATGCTTGAAGAAAGGGG	360
QY	718	TGCGCTTTCAGACAGATGATGAGAGGTGCGGCAAGATACAGCGCAGATACATGTCGAGAGA	777
Db	361	TGCGCTTTCAGACAGATGATGAGAGGTGCGGCAAGATACAGCGCAGATACATGTCGAGAGA	420
QY	778	ACAGCCCATACCTGAACCTCTAGCTGTGTTCAAGCCAGCCCAAAAAGCCATATATCAACAAATTG	837
Db	421	ACAGCCCATACCTGAACCTCTAGCTGTGTTCAAGCCAGCCCAAAAAGCCATATATCAACAAATTG	480
QY	838	TCTCAGATTTGTGTGTGTGTGTGAACCGGAGAAAGTCTATGCAATGCTACCTACTGTGCC	897
Db	481	TCTCAGATTTGTGTGTGTGTGTGAACCGGAGAAAGTCTATGCAATGCTACCTACTGTGCC	540
QY	898	CCGACAGTGAACATCAAAAGCCCTCACTAACCTGTGTGACTGTGGCCAGCCAGAGATGTGTGG	957
Db	541	CCGACAGTGAACATCAAAAGCCCTCACTAACCTGTGTGACTGTGGCCAGCCAGAGATGTGTGG	600
QY	958	TTATCATTTGATGGGCGAAGACATATTCCAGGCTTCTCCACGCTGTCTCCGCGGACCCAGA	1017
Db	601	TTATCATTTGATGGGCGAAGACATATTCCAGGCTTCTCCACGCTGTCTCCGCGGACCCAGA	660
QY	1018	TGAGGCTTCTCGAGAGTGTCTTGATGGAATTTTATCCTGTGGTGTGCTATACCGGCTCTC	1077
Db	661	TGAGGCTTCTCGAGAGTGTCTTGATGGAATTTTATCCTGTGGTGTGCTATACCGGCTCTC	720
QY	1078	TTTCATTGTAGAGTAAACTTTGTCTATGACAGACGATTATATATGACAGAAAGACAGTCCA	1137
Db	721	TTTCATTGTAGAGTAAACTTTGTCTATGACAGACGATTATATATGATGTAAGAACCAAGTCTA	780
QY	1138	AAT 1140	
Db	781	AAT 783	
RESULT 11			
AK032025			
LOCUS	AK032025	4005 bp	linear
DEFINITION	Mus musculus adult male medulla oblongata cDNA, RIKEN full-length		
	enriched library, clone:630549E01 product:estrogen related		
	receptor, beta, full insert sequence.		
ACCESSION	AK032025		
VERSION	AK032025.1	GI:26327830	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1	Carninci, P. and Hayashizaki, Y.	

TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PubMed	10349636
REFERENCE	
AUTHORS	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PubMed	11042159
REFERENCE	
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaquchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PubMed	11076861
REFERENCE	
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	5
REFERENCE	
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	6 (bases 1 to 4005)
PubMed	
REFERENCE	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, K., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001)
MEDLINE	
PubMed	
REFERENCE	
AUTHORS	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsunumi-ku, Yokohama Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ , Fax: 81-45-503-9216]
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ location/Qualifiers
FEATURES	1..4005 /organism="Mus musculus" /mol_type="mRNA" /strain="CS7BL/6J" /db_xref="FANTOM DB:G330549E01" /db_xref="MG1:2395414" /db_xref="taxon:10090"

```

/cldone="6330549E01"
/sex="male"
/tissue_type="medulla oblongata"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
143..1444
/clone="unnamed protein product; estrogen related receptor,
beta (MGI:MGI:1346832, GB|NM_011934, evidence: BLASTN,
99%, match=1315)
putative"
/codon_start=1
/protein_id="BAC27656.1"
/db_xref="GI:26327831"
/translaction="MSEDRHLGSSCSGFIKTEPSPSGIDALSHSPSGSDASGG
FGIALSTHANGLDPPMPFAGLGNPCRSKYEEDTSGIMEBSAIKCEYMNATPKRL
CLVGDIASGVHYVASCACAKAFKPTIOGNIENCGATNECEITKRRRSQCAKCP
MKCIKVCMLKRGVLDRIYRGSRQKTKRLDSENSPYLNLPISPAKPELTVISNLG
VEQKRLAMPNDIPBGDIKALTYLCELDRLGLVFLINWAKHIFGFPSTLGDQMSKL
QSAWMEILIGIVRSLPYDIDKLAAYEDYIMDEBSHSLVGLDLYRALIQIVRYKXL
KVEKEFEMILKALANSDSMYIENLEAVQCLQDLHLADQYELSGRHEPBRAGKL
LITPLRQITAKAVQHFYSYKLGQKVPMEKLFLEMLFAKY"
3984..3989
/clone="putative"
/clone="putative"
4005
/clone="putative"
polya_site
/clone="putative"
22.9%; Score 684.4; DB 11; Length 4005;
Best Local Similarity 70.0%; Pred. No. 1.8e-131;
Matches 938; Conservative 0; Mismatches 396; Indels 6; Gaps 1;

```

```

Db 705 GGCTGATTGGAGAACAGCCCTTACCTGAACTGCGCATTTTCCACCTGCTAAAGC 764
Qy 823 CATATACAAATTTGTTCTCAATTTGTTGGCTGAAACCGAGAAATCTATGCCATGC 882
Db 765 CATTACTAATGATGCTCTGCAATCTACTAGGGGTGACAGAGCAAGCTGATGCTATGC 824
Qy 883 CTGACCCCTACTGTCCTCCGACAGTACATCAAAAGCCCTCACTACATGCTGACTGGCCG 942
Db 825 CTCCCAACGATATCCCGAGGAGATATCAAGGCTTCAACCACTCTGTGAATGGCAG 884
Qy 943 ACCGAGATTTGATGATTTATCAATTTGATGAGTGGCGAAGCATATTTCCAGGCTTCCACGCTGT 1002
Db 885 ATCCGGGGCTTGTGTCTCTATCAACTGGCGCAAGCATATCCAGGCTTCCCACTCTGA 944
Qy 1003 CCTTGGCGGACCAATGAGCTTTCTGCAAGTCTTGGATGAGAAATTTGATCCTTGGTGTG 1062
Db 945 CACTTGGGAGCAAGATGAGCTTGTGCGAGATGCTCTGATGAGATTCATCTTGGGCA 1004
Qy 1063 TCGTATACCGGCTCTTTTCAATTTGAGATGAACTTGTCTATGAGACGATATATATAG 1122
Db 1005 TCGTGTACCGTCTGCTCCATACATGACAGTGTGCAATGACGAGACTATATCATAG 1064
Qy 1123 ACGAAGACAGTCCAAATTAGCAGGCTTCTTGAATCTAATATATGCTATCTGAGCTGG 1182
Db 1065 ATGAGGAACACTCTGCTGCTGTAGGGCTGTGACCTTTACCGAGCATCTGCACTGG 1124
Qy 1183 TAAAGAAATACAGAGATGAGCTGGAAGAAAGAAATTTGTAACCTCAAGCTATAG 1242
Db 1125 TGCGAGAGTACAAAACTCAAGGTAGAGAGAGAGAGTTATGATCTCCAGGCTCTGG 1184
Qy 1243 CTCTTGCATATTCAGCTCCATGACATAGAAATGTTGAAGCCGTTCAAGAGCTTCAGG 1302
Db 1185 CCTTCGCACTCACTGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1244
Qy 1303 ATGCTTCAATGAGAGCTGTGAGATTTATGAGCTGTGAGAGCTGTGAGAGCTGTG 1362
Db 1245 ACTGTGTGACAGAGGCTGTGAGAGCTATGAGCTGAGTCAAGCGCAGAGAGAGCGGGA 1304
Qy 1363 GAGCTGCAAGATGCTATGACATGCTGCTCTGAGGCAAGCTTCAACAGGCGTGGC 1422
Db 1305 GGGCGGCAAGCTGTGCTGATGAGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 1364
Qy 1423 AGCATTTCTCAACATCAAACTAGAGAGCAAGGCCAATGACAAATTTTGTGAA 1482
Db 1365 AAGCATTTCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1424
Qy 1483 TGTGGAGGCCAAGTCTGA 1502
Db 1425 TGCTGAGAGGCCAAGTGTGA 1444

```

RESULT 12
AK052256
LOCUS
DEFINITION
 Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone: D33001B12 product: estrogen related receptor, beta, full insert sequence.
ACCESSION
 AK052256
VERSION
 AK052256.1 GI:26342471
KEYWORDS
 HTC; CAP trapper.
SOURCE
 Mus musculus (house mouse)
ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 PUBMED
 99279253
 10349636
REFERENCE
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20493374
PUBMED	11042159
REFERENCE	
AUTHORS	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komoto,H., Akiyama,J., Nishii,K., Katsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,Y., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsubiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,T., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL	20530913
MEDLINE	
PUBMED	11076681
REFERENCE	
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	
AUTHORS	5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II team.
TITLE	Analysis of mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 4297) Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Horii,F., Imochani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Komo,H., Konda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ono,M., Ohsato,N., Okazaki,Y., Saito,R., Satoh,K., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sobabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takao-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physiological and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp/ URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
TITLE	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken
JOURNAL	Genomic Sciences Center and Genome Science Laboratory in RIKEN,
REFERENCE	Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
AUTHORS	Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/. location/Qualifiers
COMMENT	
FEATURES	
SOURCE	source 1. 4297 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM_DB:2330014B12" /db_xref="MG1:242118" /db_xref="taxon:10090" /clone="ID330014B12" /tissue_type="heart" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="13 days embryo"
CDS	/note="unnamed protein product; estrogen related receptor, beta [MIM:134683], GN[M 011934, evidence: BLASTN,"

```

99%, match=1315)
putative"
/codon_start=1
/protein_id="BAC34898.1"
/db_xref="GI:26342472"
/translation="MSSEDRHLGSSCGSFITKEPSSPSGIDALSHHS PGSSGDAAGG
EGIALSTHANGLS PPMFAGAGGNGNCRSYPDCTSGIMEDSAIKERYMLNIPKLT
CLVGDILASGHHGVASCEAKAPFKRTIOGNIENYCPAINBEIGITRRRKSOACAPF
MKCLIKVMPKXEVRLDRVRGRQRKTKRLDSENSVPLNLEISPAKPLTKIVSNLIG
VEODKLKAMPLEDPBGDIALTKLELAPRELVEPLINMKHILPEPSLTLGQMSLL
QSAAMELILIGIYRSLPYDPKLIAVADYIMDEBSLVLKIDLYRALIQLVRYKKIL
KSEAEEMILLKALALANSDDMYIENLEAVOKLIDILHEALQDYELSRHREPRRAGKL
LTLPLRLQRTAAKAVQHFYSVKIQGKVPMLKLELMLEAVY"
polyA_signal
4280..4285
/note="putative"
4297
polyA_site
7497
/note="putative"

```

Query Match 22.8%, Score 679.8; DB 11; Length 4297;
Best local Similarity 70.2%; Pred. No. 1.6e-130;
Matches 928; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

QY	182	GCTTCTCGCAGATNGTCAAAAGATCCGACACATGATTCGAGCTTGCTGCTTCAT	241
Db			
QY	425	GCTGCTGAACCGAATGTGCTCGAAGACAGGCACCTGGGCTCTAGTGCGCTCTTCAT	484
Db			
QY	242	CAAGACGGAACTTCCAGCCCGAGCTCCCTGACGGAAGCGGTCAACCAACGACCGCTTG	301
Db			
QY	485	CAAGACGGAACTTCCAGCCCGAGCTCCCTGACGGAAGCGGTCAACCAACGACCGCTTG	544
Db			
QY	302	TGGCTCTTCAGACGCCAGTGGAGCTACAGTTCAACCTGAATGAGCCATCAAGACGACT	361
Db			
QY	545	CGGCTCGTGGACGCCAGTGGTGGCTTGGCATTTGCCCTTGACACCCACGCCAACGGCTCT	604
Db			
QY	352	TGATCTCGGCACTCTTCACCTTCTTGCTCCATCTCTGGAGATGATGGGCTTGTCAGGA	421
Db			
QY	605	GGACTCGCGGCTTATGTTCCGACGGTGC-----GGGGCTGGAGGCAACCGTGGCGGA	658
Db			
QY	422	ACTGATATGATGATGCTCCAGACACATGTTGGAAGATCCCGACCAAGATGTGAATACAT	481
Db			
QY	659	GAGCTACGAGACCTGTACTAGTGTATCATGGAAGACCTCCGCATCAATGCGACTACAT	718
Db			
QY	482	GCTCAACTCGATGCCCAAGAGACTGTGTTTATGTGTGTGTGACATCGCTTGGGTACCA	541
Db			
QY	719	GCTTAAGCCCATCCCAAGCGCGTGTGCTGTGTGGGGGACATGTGCTCTGGCTACA	778
Db			
QY	542	CTATGGGGTAGCATATATGTGAAGCTCGGAAGGCATTTCTCAAGAGAACATTTCAAGGCA	601
Db			
QY	779	CTACGAGTGGCTCTCTCGCGAGGCTTGGAGGCGTTCCTTCAAGAGAACATTTCAAGGCA	838
Db			
QY	602	TATGAATATACGCTCCCTGCGACGAATGATGTGAATTCACAAAGCCGACGCTAATC	661
Db			
QY	839	CATCGAGTACAACTCCCGGCCACCAATGAATGTGAATTCACAAAGCCGACGCTAATC	898
Db			
QY	662	CTGCCAGGCTTGCGCTTCATGAATGTTTAAAGTGGGCAATGCTGAAGAAAGGGGTCG	721
Db			
QY	899	CTGTAGGCGCTGCGCATTCATGAATGTGCTCAAAAGTGGGAAATGTGAAGAAAGGTGG	958
Db			
QY	722	TCTTGACAGAGTACGTGGAGGTTCGGCAAGATGACAAAGCGACGATGATGTGGGAGAACAG	781
Db			
QY	959	CTTTGACCGAGTTTCAGAGAGGCGCGCAAGAGTACAGGACGCGCTGGATTCGGAACAG	1011
Db			
QY	782	CCCAATCCGAAACCTCGACGTGTTCAAGCGACGCCAAAGACATTAACAAGATGTGTC	841
Db			
QY	1019	CCCTTACCTGAACCTGCGCGATTTCCCGACCTGCTTAAAGACCATTTGACTAGATGCTTC	1077
Db			
QY	842	ACATTTGTTGATGCTGGAACCGGAAGAATCTATGCTGCTGACCCCTACTGTGCCCGA	901
Db			
QY	1079	GAACTTACTAGGGGTTGACAGGACCAAGCTGTATGCTATGCTCCCAACGATATCCCGA	1133
Db			
QY	902	CAGTGAATGAAGCCCTTCACTACCTGTGTGACTTGGCCGACGAGAGTTGTGTGTTAT	961
Db			

```

Db      1139 GGGAGATATCAAGGCCCTGACCACTCTCTGTAATTGGACAGATGGAGCTTGTTCT 1198
Qy      962 CATGGATGGGGGGAAGCATATTTCAGAGCTTCTCCACGCTGTCCCTGGGGGACAGATAG 1021
Db      1199 CATCACTGGGCGCAAGACATCCAGGCTTCCCAAGTGTGACACTGGGGACAGATAG 1258
Qy      1022 CCTTCTGAGAGTGGCTGATGAGAAATTTTATCCTTGTGCTATACCGGTCTCTTTC 1081
Db      1259 CCTGCTGAGAGTGGCTGATGAGATTTCTCATCTTGGGCATCTGTATCCGCTGCTCC 1318
Qy      1082 ATTGAGATGAATCTTCTATGACAGCATTTATTAATGACAGACACAGTCCAAAT 1141
Db      1319 ATACGATGACAAAGTGGCTGATGAGCACTATATCATGATGAGGAACACTCTCGCT 1378
Qy      1142 AGCAGGCTTCTTATCTAATTAATGCTATCTGAGCTGTAAAGAAATACAGAGAT 1201
Db      1379 GGTAGGGGTCTGACCTTTTACGAGCATCTTGAGCTGTGTGGGAGTCAAGAACT 1438
Qy      1202 GAAGCTGGAAAAAGAGATTTTGTACCCCTCAAGCTATAGCTCTTGTAAATTCAGACTC 1261
Db      1439 CAAAGTGAAGAGAGAGATTTATGATCTTCAAGGCCCTGGCCCTCGCAACTCAGATTC 1498
Qy      1262 CATGCACTAATAAATGTTGAAAGCCCTTCAAGAGCTTCAAGATGTTTACATGAAGCT 1321
Db      1499 GATGTACATTTGAGAACCTGAGGGGCTGACAGAGCTTCAAGAGCTTCTGACAGAGGCT 1558
Qy      1322 GCAGGATTTATGAAGCTGGCGACACATGAGAGACCTCTGAGCTGGCAAGATGCTGAT 1381
Db      1559 GCAGGATTTATGAAGCTGGCGACACATGAGAGACCTCTGAGCTGGCAAGATGCTGAT 1618
Qy      1382 GACACTGCTCTCTGAGGAGAGACCTTACCAAGCCGCTGAGAGATTTCTTCAACATCA 1441
Db      1619 GACCTCTGCTCTGAGGAGAGACCTTACCAAGCCGCTGAGAGATTTCTTCAACATCA 1678
Qy      1442 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1501
Db      1679 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1738
Qy      1502 A 1502
Db      1739 A 1739

```

```

RESULT 13
LOCUS   B60472
DEFINITION B60472 704 bp DNA linear GSS 21-JUN-1998
          CTT-HSP-2008K21.TRB CTT-HSP Homo sapiens genomic clone 2008K21,
          genomic survey sequence.
ACCESSION B60472
VERSION   B60472.1 GI:2615190
KEYWORDS  GSS.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 704)
          Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
          Golden,K., Berry,K., Granger,D., Sub,E., Wible,C., Shizuya,H.,
          Simon,M. and Venter,J.C.
          Use of a random BAC End Sequence Database for Sequence-Ready Map
          Building
          Unpublished (1997)
          Contact: Mark Adams
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0200
          Fax: 301 838 0208
          Email: mdadams@igf.org
          Clones are available from Research Genetics (info@resgen.com). BAC
          end search page:
          http://www.igf.org/tbdb/humgen/bac_end_search/bac_end_search.html
          Seq primer: M13 Reverse

```

```

FEATURES
Source
Class: BAC ends.
Location/Qualifiers
1..704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7040519"
/db_xref="taxon:9606"
/clone="2008K21"
/sex="Male"
/cell_type="Sperm"
/clone_id="CTT-HSP"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII"

```

```

Query Match 22.7%; Score 679; DB 28; Length 704;
Best Local Similarity 99.6%; Pred. No. 2.8e-130;
Matches 702; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

Qy      1293 AAGCTTGAGATGCTTACATGAAGGCTGAGAGATTTAGAGCTGGGACAGATGGA 1352
Db      2 AAGCTTGAGATGCTTACATGAAGGCTGAGAGATTTAGAGCTGGGACAGATGGA 61
Qy      1353 GACCCGTGCTGAGCTGGCAAGATGCTGATGACACTGACCTCTGAGGACAGACTTAC 1412
Db      62 GACCCGTGCTGAGCTGGCAAGATGCTGATGACACTGACCTCTGAGGACAGACTTAC 121
Qy      1413 AAGCCGTGACAGCTTTCTTACATCACTAAGAGGCAAAAGTCCCATGCAAACTT 1472
Db      122 AAGCCGTGACAGCTTTCTTACATCACTAAGAGGCAAAAGTCCCATGCAAACTT 181
Qy      1473 TTTTGGAAATGTTGAGAGGCAAGGCTGACTAAGAGCTCCGAGCTTCCATCTTC 1532
Db      182 TTTTGGAAATGTTGAGAGGCAAGGCTGACTAAGAGCTCCGAGCTTCCATCTTC 241
Qy      1533 ATGTTGAAAAAGGAAAAATAAACCAAGATGATGTCGAAAGAACTTAGATTAGTTAA 1592
Db      242 ATGTTGAAAAAGGAAAAATAAACCAAGATGATGTCGAAAGAACTTAGATTAGTTAA 301
Qy      1593 CAACATCAAAATCAACAGACTGACATGATATTATAGAGCAAGACATAGAGAGCTT 1652
Db      302 CAACATCAAAATCAACAGACTGACATGATATTATAGAGCAAGACATAGAGAGCTT 361
Qy      1653 CAGATTCCTGCAATGAGTTCCTGATGAG-TTCTTCTACTTCTGATCATCTTCTCTCT 1711
Db      362 CAGATTCCTGCAATGAGTTCCTGATGAG-TTCTTCTACTTCTGATCATCTTCTCTCT 421
Qy      1712 CTTTCTCCCAATTTCTCTTCTCTTATTTTCTCTTCTCTTCTCTTCTCTTCTCTTCT 1771
Db      422 CTTTCTCCCAATTTCTCTTCTCTTATTTTCTCTTCTCTTCTCTTCTCTTCTCTTCT 481
Qy      1772 ATTCTTCTGCTCTTCTATTCCTAGTTCCTTCTCTTATTTTCTCTTCTCTTCTCT 1831
Db      482 ATTCTTCTGCTCTTCTATTCCTAGTTCCTTCTCTTATTTTCTCTTCTCTTCTCT 541
Qy      1832 CCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1891
Db      542 CCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 601
Qy      1892 TAAATTTGAATAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1951
Db      602 TAAATTTGAATAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 659
Qy      1952 TTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1996
Db      660 TTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 704

```

```

RESULT 14
LOCUS   AY409736
DEFINITION AY409736 1296 bp DNA linear GSS 16-DEC-2003
          Mus musculus ESRB gene, VIRUAL TRANSCRIPT, partial sequence,
          genomic survey sequence.

```

ACCESSION AY409736
 VERSION AY409736.1 GI:39765704
 KEYWORDS GSS.
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1296)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanendau,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J., Adams,M.D. and Cargill,M.
 TITLE Interfering nonneutral evolution from human-chimp-mouse orthologous gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003).
 REFERENCE 2 (bases 1 to 1296)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanendau,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J., Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
 FEATURES
 source location/Qualifiers
 1..1296
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>1296
 /gene="BSRNB"
 /locus_tag="HCM3656"
 ORIGIN
 Query Match 22.5%; Score 670.4; DB 29; Length 1296;
 Best Local Similarity 70.3%; Pred. No. 1.6e-128;
 Matches 915; Conservative 0; Mismatches 381; Indels 6; Gaps 1;
 QY 195 ATGTCAACAAAGATGACACATTTGATTCAGCTGTCCTCTTCAAGACGAACT 254
 DB 1 ATGTCTCGAAGACAGGACCTTGAGCTTATGTTGGGCTCTTCTTCAAGACGAACT 60
 QY 255 TCCAGCCACGCTCCCTGACGACAGCTCAACACACACAGCCCTGAGCTTTCAGC 314
 DB 61 TCCAGCCCTCCCTCGGACATGATGCTCAGACACACAGCCCTGAGCTTTCAGC 120
 QY 315 GCCAGTGGAGACTACGTTCAACCATGATGCCCATCAGAAAGGACTTGACTGCCACT 374
 DB 121 GCCAGTGGAGCTTGGCATTTGCCCTGAGCACACGCTGAGCTTGACTGCCACT 180
 QY 375 CTCTACCTTCTGCTCTTATCCTCGGAGGATGAGGCGCTGTCAAGAACTGATGAC 434
 DB 181 ATGTTGACAGGTGC-----GGGCTGGAGGCAACCCGCGCAAGACTACAGAGAC 234
 QY 435 TGCTCCAGACCAATTTGTAAGATCCCAAGACCAAGTGTGATATACATGCTCAACTGATG 494
 DB 235 TGTACTAGTGTATCATGAGGACTCCGACCATCAATGAGATACATGCTTAAGCATC 294
 QY 495 CCCAAGAGACTGTGTTAGTGTGTGTGATGCTTCTGGTACCACTATGGGCTAGCA 554
 DB 295 CCCAAGAGCTGTGCTCGTGTGCGGAGCATTTGCCCTGTGCTACCACTACGAGTGGCC 354
 QY 555 TCAATGGAAGCCGCAAGGACTTCTTAAGAGACCAATTAAGGCAATATGAAATCAGC 614
 DB 355 TCTGTGAGGCTTGAAGGCGTCTTCAAGAGACCAATTAAGGCAATATGAGTACAC 414
 QY 615 TGCCCTGCCACGATGATGTAATCAAAAGGACGAGAGTAAATCCTCGCAAGCTTGC 674
 DB 415 TGCCCGGCCACGATGATGTAATCAAAAGGAGGAGGCAAGTCTCTGCAAGGCTCGC 474
 QY 675 CGCTTCATGAAGTGTAAAAGTGGGCACTGCTGAAGAAAGAGGGTGGCTTTCAGACAGTA 734

DB 475 CGATTTCATGAATCCCTCAAGTGGGATGCTGAAGAGAGTGGCGCTTGACCAAGTT 534
 QY 735 CGTGAAGTGGGACAAAGTACAGGACGATGATGCGGAGAACAGCCCATCTGTAAC 794
 DB 535 CGAGAGGCGCCGCAAGATACAGGACGCTGATTTGGAGAACAGCCCATCTGTAAC 594
 QY 795 CCTCAGCTGTTCCAGCCGCAAAAGCCATTAACAAGATTTGTCATTTGTTGGG 854
 DB 595 CTGCGGATTTCCCACTGCTAAGAAAGCCATTAAGATGTTGTCATTTGTTGGG 654
 QY 855 GCTGAACCGGAAGATCTATGCACTGCTGACCTGATCCCTGTCCTCCAGAGTACATCAA 914
 DB 655 GTTAGCGAGACAACTATGCTATGCTATGCTCCCAAGATATCCCGAGGAGATATCAA 714
 QY 915 GCCCTCACTACATGTTGACTTGGCCGACGAGAGTTGGTGTATATGATGGAGCG 974
 DB 715 GCCCTGACCACTCTCTGTAATTTGGCAATGGGAGCTTGTGTTCTCTCACTACCTGGGCG 774
 QY 975 AAGCATTTCCAGGTTTCCACGCTGTCCTGGGAGCCAGATGAGGCTTGGCAGAGT 1034
 DB 775 AAGCATTTCCAGGTTTCCACGCTGTCCTGGGAGCCAGATGAGGCTTGGCAGAGT 834
 QY 1035 GCTTGATGAAATTTGATCCTTGTGTCGTAACCGGTCCTTTTCATTTGAGATGA 1094
 DB 835 GCTTGATGAGATTTCTATCTTGGGCACTGCTGTAACCGGTCCTTCCATACATGACAG 894
 QY 1095 CTGTCTATGACAGATTAATTAATGACAGAACAGCTCCAAATTTAGCAGGCTTCTT 1154
 DB 895 CTGCAATGCGCCGAGGACTATATCATGATGAGAGAACCTCTGCGCTGTAGGCTGCTG 954
 QY 1155 GATCTAAATATGCTATCTCTGAGCTGTGAAGAAATTAAGAGCATGAAGCTGGAATA 1214
 DB 955 GACCTTTACCGAGCATCTCTGAGCTGTGCGCAGGTACAAAGAACTCAAGTTAGAGAG 1014
 QY 1215 GAAGATTTGTACCTCAACCTTAAGCTTGTCTTCAATTCAGACTCCATGACATAGAA 1274
 DB 1015 GAAGATTTGTATCTCTCAAGGCTGCGCCCTCGCAACTCATGATGATGATGAG 1074
 QY 1275 GATGTTGAAGCGGTTGAGAGCTTCAAGATGCTTACATGAAGGCTGAGAGTTATGA 1334
 DB 1075 AACCTGAGGCGGTGAGAGAGCTCCAGACCTGCTGACAGAGCGCTGAGAGCTATGAG 1134
 QY 1335 GCTGCGACCAATGGAAGAACCTGCTGAGCTGCGAAGATGATGACATGACAGCACTC 1394
 DB 1135 CTGAGTCAGCGCAACAGAGAGCGCGAGAGCGCGAAGCTGCTGAGAGCTGCGCCCTG 1194
 QY 1395 CTGAGGCAAGACCTTCAACAGGCGGTCAGCACTTTCTACAGATCAAACTAGAGGCAAA 1454
 DB 1195 CTGAGGCAAGACCGCCCAAGCGCTGCAACTTCTACAGTGTGAATGTCAGGGCAAG 1254
 QY 1455 GTCCCAATGACAAACTTTTGTGAAGATGTTGAGGCCAAG 1496
 DB 1255 GTGCCCATGACAAACTTTCTCTGAGATGCTGAGAGGCCAAG 1296
 RESULT 15
 CB989379 780 bp mRNA linear EST 01-MAY-2003
 LOCUS AGENCOURT 13902955 NIH-MGC 147 Homo sapiens cDNA clone
 DEFINITION IMAGE:30345143 5', mRNA sequence.
 ACCESSION CB989379
 VERSION CB989379.1 GI:30283899
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 780)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: c9apbs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
 Plate: NDAM382 row: 1 column: 24
 High quality sequence stop: 606.

FEATURES

Location/Qualifiers
 1..780
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30345143"
 /tissue_type="Human Placenta"
 /lab_host="DH10B Tona"
 /clone_id="NIH MGC 147"
 /note="Organ: placenta; Vector: pbluescriptR; Site 1:
 all-XhoI; Site 2: BamH; Oligo-dT primed using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average
 insert size 2.3 kb and normalized to ROT 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: This is
 a NIH_MGC library."

ORIGIN

Query Match 22.4%; Score 667.6; DB 14; Length 780;

Best Local Similarity 98.4%; Pred. No. 6.4e-128;

Matches 685; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

```

QY 17 GTGGAATTGGCTGTGCTAGAGAAATTTGTTATGCACTGCTCTGTCAAG 76
DB 85 GTGCTTAGGGCTGTGCTAGAGAAATTTGTTATGCACTGCTCTGTCAAG 144
QY 77 AAACCTTGATTATAGCTGGGGTGCACAATATATGTTGCCGTCGCATGATGGAT 136
DB 145 AAATTTGATTATAGCTGGGGTGCACAATATATGTTGCCGTCGCATGATGGAT 204
QY 137 AGAATTTGCTTCTGTAATCTTTTCTGCTAGAGAAAGCTTCTCTGCAAGAT 196
DB 205 AGAATTTGCTTCTGTAATCTTTTCTGCTAGAGAAAGCTTCTCTGCAAGAT 264
QY 197 GTCAACAAGATGACATTTGATTCAGTGTTCCTTCTCATCAAGACGAAACCTTC 256
DB 265 GTCAACAAGATGACATTTGATTCAGTGTTCCTTCTCATCAAGACGAAACCTTC 324
QY 257 CAGCCCAAGCTCCCTGAGAGACAGCTCAACCAACAGCCCTGTGGCTCTTCAAGGC 316
DB 325 CAGCCCAAGCTCCCTGAGAGACAGCTCAACCAACAGCCCTGTGGCTCTTCAAGGC 384
QY 317 CAGTGGAGGCTACAGTTCACATGAATGGCCATCAGAACGCACTTGACTGCCACCTCT 376
DB 385 CAGTGGAGGCTACAGTTCACATGAATGGCCATCAGAACGCACTTGACTGCCACCTCT 444
QY 377 CTACCCCTTGTGCTCTAATCCCGGAGGTAATGGGCTGTCAAGAACTGTATGATGACTG 436
DB 445 CTACCCCTTGTGCTCTAATCCCGGAGGTAATGGGCTGTCAAGAACTGTATGATGACTG 504
QY 437 CTCAGACCATGTTGAAGATCCCGAGACCAAGTGTGAATACATGCTCAACTGATGCC 496
DB 505 CTCAGACCATGTTGAAGATCCCGAGACCAAGTGTGAATACATGCTCAACTGATGCC 564
QY 497 CAAGAGACTGTGTTAGTGTGTGTCATGCTTCTGGGTACCATATGGGGTAGCATC 556
DB 565 CAAGAGACTGTGTTAGTGTGTGTCATGCTTCTGGGTACCATATGGGGTAGCATC 624
QY 557 ATGTGAAGCTGCAAGGCATTTCTCAAGAGACATTCAGGCAATATAGATACAGCTG 616

```

```

DB 625 ATGTGAAGCTGCAAGGCATTTCTCAAGAGACATTCAGGCAATATAGATACAGCTG 684
QY 617 CCTTCCACGATGATGATGAAATCAAAAGCGCAGAGCTAATCTCTGCC--AGGCTTGC 674
DB 685 CCTTCCACGATGATGATGAAATCAAAAGCGCAGAGCTAATCTCTGCCAGGCTTGC 744
QY 675 CGCTTCATGAAGTGTTTAAAGTGGGCATGCTGAAA 710
DB 745 CGCTTCATGAAGTGTTTAAAGTGGGCATGCTGAAA 780

```

Search completed: August 10, 2004, 04:55:27
 Job time: 7350 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2004, 23:15:03 ; Search time 208 Seconds
(without alignments)
7964.079 Million cell updates/sec

Title: US-10-054-841-3
Perfect score: 2985
Sequence: 1 gcggcgccgacagtggtggtg.....actggcgccgacgcttgat 2985

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 segs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seg:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seg:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seg:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seg:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seg:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seg:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1559	52.2	1615	US-09-040-508-1	Sequence 1, Appl1
2	1559	52.2	1615	US-09-500-654-1	Sequence 1, Appl1
3	443.2	14.8	536	US-09-040-508-3	Sequence 3, Appl1
4	443.2	14.8	536	US-09-500-654-3	Sequence 3, Appl1
5	421.4	14.1	2539	US-09-620-312D-454	Sequence 454, Appl
6	162.6	5.4	7218	US-08-232-463-14	Sequence 14, Appl
7	136.6	4.6	4963	US-08-076-726-16	Sequence 16, Appl
8	136.6	4.6	4963	US-08-260-452-9	Sequence 9, Appl1
9	136.6	4.6	4963	US-08-481-970-9	Sequence 9, Appl1
10	136.6	4.6	4963	US-08-897-719-9	Sequence 9, Appl1
11	136.6	4.6	4963	US-09-163-269-9	Sequence 9, Appl1
12	136.6	4.6	6450	US-09-041-886-34	Sequence 34, Appl
13	136.6	4.6	6450	US-08-453-988-3	Sequence 3, Appl1
14	136.6	4.6	6450	US-10-027-983-3	Sequence 3, Appl1
15	131.6	4.4	472	US-09-833-381-269	Sequence 269, App
16	131.6	4.4	472	US-09-833-381-270	Sequence 270, App
17	125.8	4.2	1757	US-08-216-592A-3	Sequence 3, Appl1
18	125.8	4.2	2130	US-07-952-800-1	Sequence 1, Appl1
19	125.8	4.2	2204	US-07-952-800-3	Sequence 3, Appl1
20	125.8	4.2	2219	US-08-336-408B-7	Sequence 7, Appl1
21	125.8	4.2	2219	PCT-US91-00399-7	Sequence 7, Appl1
22	124.6	4.2	2928	US-08-095-728B-3	Sequence 3, Appl1
23	124.6	4.2	2928	PCT-US92-02320A-3	Sequence 3, Appl1
24	124.6	4.2	2940	US-08-592-383-1	Sequence 1, Appl1
25	124.6	4.2	3036	US-08-106-691B-52	Sequence 52, Appl
26	124.6	4.2	3036	US-08-095-728B-1	Sequence 1, Appl1
27	124.6	4.2	3036	PCT-US92-02320A-1	Sequence 1, Appl1

28	124.6	4.2	3511	US-08-892-747-13	Sequence 13, Appl
29	124.2	4.2	2285	US-08-216-592A-1	Sequence 1, Appl1
30	123.6	4.1	1576	5260432-1	Patent No. 5260432
31	123	4.1	2940	5171671-1	Patent No. 5171671
32	122.6	4.1	1594	US-09-016-434-1059	Sequence 1059, Ap
33	120.4	4.0	1883	US-08-216-592A-5	Sequence 5, Appl1
34	120.4	4.0	2095	US-08-336-408B-3	Sequence 3, Appl1
35	120.4	4.0	2095	PCT-US91-00399-3	Sequence 3, Appl1
36	118.2	4.0	1866	US-08-336-408B-1	Sequence 1, Appl1
37	118.2	4.0	1866	PCT-US91-00399-1	Sequence 1, Appl1
38	113.8	3.8	1251	US-09-608-088-2	Sequence 2, Appl1
39	113.8	3.8	1251	US-09-608-088-20	Sequence 20, Appl
40	113.8	3.8	1434	US-08-836-620A-4	Sequence 4, Appl1
41	113.8	3.8	1460	US-08-836-620A-4	Sequence 4, Appl1
42	113.8	3.8	1647	US-09-139-617-2	Sequence 2, Appl1
43	113.8	3.8	1647	US-09-561-741A-2	Sequence 2, Appl1
44	113.8	3.8	1647	US-09-558-795-2	Sequence 2, Appl1
45	113.8	3.8	1898	US-09-608-088-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-040-508-1
; Sequence 1, Application US/09040508
; Patent No. 6069239
; GENERAL INFORMATION:
; APPLICANT: Mathias, Stephen L.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSES: RATHER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 17-MAR-1998
; APPLICATION NUMBER: US/09/040,508
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9705451.4
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: GB 9803289.9
; FILING DATE: 16-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-30008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-09-040-508-1
Query Match 52.2%; Score 1559; DB 3; Length 1615;
Best local similarity 99.7%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 17 GTGGAATTGGCTTGTCACTAGAGAACATTGTGTATTTGCACTGTGCTGTCAAG 76
 Db 49 GTGCTTAATAGGTTGTCTACTAGAGAACATTGTGTATTTGCACTGTGCTGTCAAG 108
 QY 77 AAACCTTATTTATTTAGCTGGGGGTGCAAAATATATGTTTCCGGTGCACATGATTCGCT 136
 Db 109 AAACCTTATTTATTTAGCTGGGGGTGCAAAATATATGTTTCCGGTGCACATGATTCGCT 168
 QY 137 AGAATTGGCTTGTCACTAGAGAACATTGTGTATTTGCACTGTGCTGTCAAG 196
 Db 169 AGAATTGGCTTGTCACTAGAGAACATTGTGTATTTGCACTGTGCTGTCAAG 228
 QY 197 GTCAAAACAAGATGACACATTGATTCAGCTGTGCTGTCACTAGAGAACATTG 256
 Db 229 GTCAAAACAAGATGACACATTGATTCAGCTGTGCTGTCACTAGAGAACATTG 288
 QY 257 CAGCCCAAGCTTCTGAGAGACAGCTGTCAACCAAGCCCTGTGCTGTCAAGAGC 316
 Db 289 CAGCCCAAGCTTCTGAGAGACAGCTGTCAACCAAGCCCTGTGCTGTCAAGAGC 348
 QY 317 CAGTGGAGAGCTACAGTCAACATGAAATGGCCATCAGAGAGGACCTTGACCTGCA 376
 Db 349 CAGTGGAGAGCTACAGTCAACATGAAATGGCCATCAGAGAGGACCTTGACCTGCA 408
 QY 377 CTACCTCTTGTCTCTATCTCTGAGAGATGAGGCTGTCAAGAACTGTATGATGACTG 436
 Db 409 CTACCTCTTGTCTCTATCTCTGAGAGATGAGGCTGTCAAGAACTGTATGATGACTG 468
 QY 437 CTCCAGACCACTTTGTGAAGATCCCCAGACCAAGTGTGAATACATGCTCAACTGATGCC 496
 Db 469 CTCCAGACCACTTTGTGAAGATCCCCAGACCAAGTGTGAATACATGCTCAACTGATGCC 528
 QY 497 CAAGAGACTGTGTAGTGTGTGATGATGCTGTGAGATGAGGAGTGTGAGATG 556
 Db 529 CAAGAGACTGTGTAGTGTGTGATGATGCTGTGAGATGAGGAGTGTGAGATG 588
 QY 557 ATGTGAAGCTTGAAGGATCTTCAAGAGACATTCAGAGCAATATGAAATACAGCTG 616
 Db 589 ATGTGAAGCTTGAAGGATCTTCAAGAGACATTCAGAGCAATATGAAATACAGCTG 648
 QY 617 CCCGSCCAAGATGAATGTGAATCAAAAGCCAGACGTAATCTTCCAGGCTTGGCG 676
 Db 649 CCCGSCCAAGATGAATGTGAATCAAAAGCCAGACGTAATCTTCCAGGCTTGGCG 708
 QY 677 CTTTCATGAAGTGTAAAAGTGGGCTGTGAAGAAAGGGGTGGCTGTGACAGATGAG 736
 Db 709 CTTTCATGAAGTGTAAAAGTGGGCTGTGAAGAAAGGGGTGGCTGTGACAGATGAG 768
 QY 737 TGGAGGTGCGAGAAATGACAAAGCCAGAGATGATGCGAGAAACAGCCATACCTGAAG 796
 Db 769 TGGAGGTGCGAGAAATGACAAAGCCAGAGATGATGCGAGAAACAGCCATACCTGAAG 828
 QY 797 TCGAGTGGTTCAGGCAAGCAAAAGCCATATTAACAATGTCTCAATTTGTGTGGGC 856
 Db 829 TCGAGTGGTTCAGGCAAGCAAAAGCCATATTAACAATGTCTCAATTTGTGTGGGC 888
 QY 857 TGAACCGAGAGATCTATGCACTGCTGACCCCTACTGCTCCCGACAGTACATCAAGC 916
 Db 889 TGAACCGAGAGATCTATGCACTGCTGACCCCTACTGCTCCCGACAGTACATCAAGC 948
 QY 917 CTTTCATGAAGTGTAAAAGTGGGCTGTGAAGAAAGGGGTGGCTGTGACAGATGAG 976
 Db 949 CTTTCATGAAGTGTAAAAGTGGGCTGTGAAGAAAGGGGTGGCTGTGACAGATGAG 1008
 QY 977 GCATATTCAGAGCTTCCACAGCTGTGCTGGGCGACACATGAGCTTCTGCAAGATGC 1036
 Db 1009 GCATATTCAGAGCTTCCACAGCTGTGCTGGGCGACACATGAGCTTCTGCAAGATGC 1068
 QY 1037 TTGGATGGAATTTTGAATCCTTGGTGTGATATACCGGCTCTCTTCATTTGAGATGACT 1096
 Db 1069 TTGGATGGAATTTTGAATCCTTGGTGTGATATACCGGCTCTCTTCATTTGAGATGACT 1128
 QY 1097 TGTCTATGACAGATTTATATATGACGAAAGCAAGTCCAAATTAAGCAGGCTTCTTGA 1156

Db 1129 TGTCTATGACAGATTTATATATGACGAAAGCAAGTCCAAATTAAGAGGCTTCTTGA 1188
 QY 1157 TCTAATATATGCTATCTGACAGCTGTGTAAGAAATATACAGACATGAGCTGAAAAAGA 1216
 Db 1189 TCTAATATATGCTATCTGACAGCTGTGTAAGAAATATACAGACATGAGCTGAAAAAGA 1248
 QY 1217 AGAATTTGTACCTCTCAAGCTATAGCTTGTCTTAATTCAGACTTCATGACATGAAAGA 1276
 Db 1249 AGAATTTGTACCTCTCAAGCTATAGCTTGTCTTAATTCAGACTTCATGACATGAAAGA 1308
 QY 1277 TGTGAAGCCGCTCAGAAAGCTTCAAGATGCTTTCATATGAAAGCGTGCAGATTAAGAGC 1336
 Db 1309 TGTGAAGCCGCTCAGAAAGCTTCAAGATGCTTTCATATGAAAGCGTGCAGATTAAGAGC 1368
 QY 1337 TGGCCAGACATGAAAGACCTCTGTGAGCTGAGCAAGATGTGTATGACACTGCCACTCT 1396
 Db 1369 TGGCCAGACATGAAAGACCTCTGTGAGCTGAGCAAGATGTGTATGACACTGCCACTCT 1428
 QY 1397 GAGGCAAGCTCTTACCAAGGCGCTGACATTTCTTCAACATCAACTGAAAGCAAGT 1456
 Db 1429 GAGGCAAGCTCTTACCAAGGCGCTGACATTTCTTCAACATCAACTGAAAGCAAGT 1488
 QY 1457 CCCATGCACAAACTTTTGTGAAGATGTTGAGAGCCCAAGGCTGACTAAAGCTCCCTG 1516
 Db 1489 CCCATGCACAAACTTTTGTGAAGATGTTGAGAGCCCAAGGCTGACTAAAGCTCCCTG 1548
 QY 1517 GGGCTTCCATCTCTTATGTTGAAAAAGGAAAAATAAACCCAGAGTATGTGAGAA 1576
 Db 1549 GGGCTTCCATCTCTTATGTTGAAAAAGGAAAAATAAACCCAGAGTATGTGAGAA 1608
 QY 1577 CTTAGAG 1583
 Db 1609 CTTAGAG 1615

RESULT 2
 US-09-500-654-1
 : Sequence 1, Application US/09500654
 : Patent No. 6359116
 : GENERAL INFORMATION:
 : APPLICANT: Stephen L. Mathias
 : TITLE OF INVENTION: No. 6359116el Compounds
 : FILE REFERENCE: GP-30008-D1
 : CURRENT APPLICATION NUMBER: US/09/500,654
 : EARLIER FILING DATE: 2000-02-09
 : EARLIER APPLICATION NUMBER: 9705451.4
 : EARLIER FILING DATE: 1997-03-17
 : EARLIER APPLICATION NUMBER: 9803289.9
 : NUMBER OF SEQ ID NOS: 4
 : SOFTWARE: PASCSEQ for Windows Version 3.0
 : SEQ ID NO 1
 : LENGTH: 1615
 : TYPE: DNA
 : ORGANISM: HOMO SAPIENS
 : US-09-500-654-1

Query Match 52.2%; Score 1559; DB 4; Length 1615;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1562; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 17 GTGGAATTGGCTTGTCACTAGAGAACATTGTGTATTTGCACTGTGCTGTCAAG 76
 Db 49 GTGCTTAATAGGTTGTCTACTAGAGAACATTGTGTATTTGCACTGTGCTGTCAAG 108
 QY 77 AAACCTTATTTATTTAGCTGGGGGTGCAAAATATATGTTTCCGGTGCACATGATTCGCT 136
 Db 109 AAACCTTATTTATTTAGCTGGGGGTGCAAAATATATGTTTCCGGTGCACATGATTCGCT 168
 QY 137 AGAATTGGCTTGTCACTAGAGAACATTGTGTATTTGCACTGTGCTGTCAAG 196
 Db 169 AGAATTGGCTTGTCACTAGAGAACATTGTGTATTTGCACTGTGCTGTCAAG 228

Matches 497; Conservative 0; Mismatches 14; Indels 6; Gaps 4;

QY	17	GTGGAAATTCGGCTTGTCTACATAGAGAAACATTTGTGTAAATTGCACTGTGTCTCTGTCAAG	76
Db	20	GTGCTTAATAGCGTTGTCTACATAGAGAAACATTTGTGTAAATTGCACTGTGTCTCTGTCAAG	79
QY	77	AAACTTTGATTTATATAGCTGGGGGTGCACAATAA--TGGTATCGCGTGCACATGATTC--G	134
Db	80	NAACCTTGTATTATAGCTGGGTGCACAATAATATGGTTGCCGGTGCACATGGATTCCG	139
QY	135	GTAGAACTTTGGCTTCTCTGAATCTTTTTTCCCTGCACTACGAGGAAG--CTTCTCTGC	191
Db	140	GAAAGAACTTTACTCTCTGAATCTTTTTTCCCTGCACTACGAGGAAGAACTTTCTCTGC	199
QY	192	AGAATGT--CAAACAAGAATCGACACATTTGATTCACAGCTGTTGGCTCTTATCAAGACGA	250
Db	200	AGAATGTCCAAACAAAGAATCGACACATTTGATTCACAGCTGTTGGCTCTTATCAAGACGA	259
QY	251	ACCTTCCAGCCCAAGCTCTCTTACGGAACAGGCTCAACACACAGCTCTGGTGGCTCTTC	310
Db	260	ACCTTCCAGCCCAAGCTCTCTTACGGAACAGGCTCAACACACAGCTCTGGTGGCTCTTC	319
QY	311	AGAGCCAGTGGGAGCTACAGTTCAACCATGAATGGCCATCGAAGCGACTTGACTCGGC	370
Db	320	AGAGCCAGTGGGAGCTACAGTTCAACCATGAATGGCCATCGAAGCGACTTGACTCGGC	379
QY	371	ACCTTCTTAACTTTCTGTCTCTATCTCTGGAGGTATGTGGGCTGTCAAGAAACTGTATGA	430
Db	380	ACCTTCTTAACTTTCTGTCTCTATCTCTGGAGGTATGTGGGCTGTCAAGAAACTGTATGA	439
QY	431	TGACTGTCCAGCACATTTGTAAGATCCCCAGACCAAGTGTGAATATGATGTCACATC	490
Db	440	TGACTGTCCAGCACATTTGTAAGATCCCCAGACCAAGTGTGAATATGATGTCACATC	499
QY	491	GATGCCCAAGAGACTGTGTATTAGTGTGTGTGATCAATC	527
Db	500	GATGCCCAAGAGACTGTGTATTAGTGTGTGTGATCAATC	536

RESULT 4
US-09-500-654-3

```

? Sequence 3, Application US/09500654
? Patent No. 6359116
?
? GENERAL INFORMATION:
? APPLICANT: Stephen L. Mathias
? TITLE OF INVENTION: No. 6359116e1 Compounds
? FILE REFERENCE: GP-30008-0D1
? CURRENT APPLICATION NUMBER: US/09/500,654
? CURRENT FILING DATE: 2000-02-09
? EARLIER APPLICATION NUMBER: 9705451.4
? EARLIER FILING DATE: 1997-03-17
? EARLIER APPLICATION NUMBER: 9803289.9
? EARLIER FILING DATE: 1998-02-16
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 3
?
? LENGTH: 536
?
? TYPE: DNA
?
? ORGANISM: HOMO SAPIENS
?
? FEATURE:
?
? NAME/KEY: UNSTUR
?
? LOCATION: (80)
?
? US-09-500-654-3

```

Query Match	14.8%	Score 443.2;	DB 4;	Length 536;
Best Local Similarity	96.1%	Pred. No. 3.7e-111;		
Matches 497; Conservative	0;	Mismatches 14;	Indels 6;	Gaps 4

QY 17 GTGAAATTCGGCTGTCACTAGAGAAACATTGTGTATTTGCACTGTGCTGTCAAG 76
|||
Db 20 GTGCTATAGCTGTCACTAGAGAAACATTGTGTATTTGCACTGTGCTGTCAAG 79
QY 77 AAACCTTGATTTATAGCTGGGGGTGCACAAATTA-TGTGTGCGGCTGGACATGATTC-G 134

Db 80 NAACTTTGATTATAGCTGGGGTGCACAAATAATTGTTGCCGGTCCGACATGGATTCCG 139

QY	133	GTGACCTTTGGCTTCTCTTAATCTTTTCCCTGCACATACGAGAAAG--CTTCTCTGC	191
Db	140	GAAGAAGCTTAACTCTTCTTAATCTTTTCCCTGCACATACGAGGAAAGAGCTTCTCTGC	199
QY	192	AGAATGT-CAACAAGATCGACACATGATTCAGCTGTTGGTCTCATCAGACGSA	250
Db	200	AGATGTCCAAACAAAGATCGACACATGATTCAGCTGTTGGTCTCTTCATCAAGACGSA	258
QY	251	ACCTTCAGCCAGCCTCCCTGACGAGACAGGCTCAACCAACAGCCCTGGTGCTCTTC	310
Db	260	ACCTTCAGCCAGCCTCCCTGACGAGACAGGCTCAACCAACAGCCCTGGTGCTCTTC	319
QY	311	AGAGCGACATGGAGAGCTACAGTTCAACATGAATGGCCATACAGACGACTTGACTGCC	370
Db	320	AGAGCGACATGGAGAGCTACAGTTCAACATGAATGGCCATACAGACGACTTGACTGCC	379
QY	371	ACCTCTCAACCTTCTGCTCTATCTCTGGAGATGAGGGGCTCTCAGSAAACGTATGA	430
Db	380	ACCTCTCAACCTTCTGCTCTATCTCTGGAGATGAGGGGCTCTCAGSAAACGTATGA	439
QY	431	TGACTGCTCCAGACCAATTTGTAAGATCCCAAGCAAGTGAATATACGTCAATC	490
Db	440	TGACTGCTCCAGACCAATTTGTAAGATCCCAAGCAAGTGAATATACATGCTCAATC	499
QY	491	GATGCCAAGACATGTGTTTAGTGTTGGTGAGATC	527
Db	500	GATGCCAAGACATGTGTTTAGTGTTGGTGAGATC	536

RESULT 5
US-09-620-312D-454

Sequence 455, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Dimañac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: Polypeptides
FILE REFERENCE: 784CIP2H
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 454
LENGTH: 2539

APPLICATION NUMBER: 08/076,327
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4963 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
IMMEDIATE SOURCE:
CLONE: PUHD BGR4
US-08-260-452-9

Query Match 4.6%; Score 136.6; DB 1; Length 4963;
Best Local Similarity 52.3%; Pred. No. 7,8e-27;
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

```
QY 507 TGTGTTAGTGTGTGTCATCGCTTCTGGTACCATATGSGGTAGCATCATGTGAAGC 566
DB 1018 TGTGCAAGTGTGAATGATGATGCTTCAAGCTACCATATGAGTCTGGTCTGTGAGGCG 1077
QY 567 TGCAGGCAATTTCTTAAGAGCAATTCAAGGCATATATGATATACAGTCCGCTCCACG 626
DB 1078 TGCAGGCTTCTTCAAGAGATATCAAGACATTAACGACTATATGTCTCAGCCAC 1137
QY 627 AATGATGTGAATCAAAAGCGACGTAATCTGTCCAGGCTTGCCTTCATGAAG 686
DB 1138 AACGATGACCATGATATATAAACAAGAGAGAGAGTGCACGCTCCGCGCTCCGCAAA 1197
QY 687 TGTGTTAAAGTGGGCAATGCTGAAGAAGGGGTGCGCTTTGACAGATCATGTGAAGTGG 746
DB 1198 TGTCTAGAAAGTGGGATGATGAAGGTGGGATACGAAAGACGGAAGAGAGAGATG 1257
QY 747 CAGAGATACAAAGCGAGATAGATGGGAGAA----- 778
DB 1258 TTGAAACACAAAGCGCCAGAGAGATGATGGGAGGCGAGGCGTGAAGTGGGCTGCTGA 1317
QY 779 -----CAGCCATACCTGAACCTCAGCTGCTTCAAGCCACCAAAA----- 820
DB 1318 GACATGAGAGCTGCAACCTTGGCCAAAGCCGCTCATATGATCAAAAGCTCTAAGAGAAC 1377
QY 821 -----GCCATATAACAAGATTTGTCTCAATTTGTTGTGCTGAACG 863
DB 1378 AGGCTGGCTTGTCCCTGACGGCCGACAGATGTCATGCTTGTGATGCTGAGCC 1437
QY 864 GAGAAATCTATGCCATGCTGACCTTACTGTCCCGACAGTACATCAAAAGCCCTCACT 923
DB 1438 CCCATCTCTATTTCCAGATATGATCTTACCAAGCCCTTCAAGTAACTTCGATGATGGG 1497
QY 924 ACACGTGTGACTTGGCCGACGAGAGTGTGTTATCATTTGATGAGGAGCATAT 983
DB 1498 TTAAGTCAACACTGGCAGAGAGAGAGTGTGTTCAATGATCAACTGGGAGAGAGAGG 1557
QY 984 CCAGGCTTCCACGCTGTCCCTGGGGAACAGATGAGCCTTTCAGAGTGTGAGTGA 1043
DB 1558 CCAGGCTTGTGATTTGACCCCTCCAWGATCAGGTCACCTCTAAGATGTGCTGGCTA 1617
QY 1044 GAAATTTTATCTTGTGTGCTGATATACGGTCTCTTCAATTTGATGAGATGAACCTTGC 1103
DB 1618 GAAATCTGATGATGTGCTGCTGCTGCTTCCATGAGACCCAGTGAAGTACTGTT 1677
QY 1104 GCAGAGATATATATGAGAGAGAACCAAGTCCAAAT---TAGAGGCTTCTTGTGATCTA 1160
DB 1678 GCTCTTACTTGTGCTTGTGACAGAGAACCAAGGAAATGTGTAAAGGCAATGTGAGATC 1737
```

```
QY 1161 AATAATGCTATCCCTGCAAGCTGTAAAGAAATACAGAGACATGAGAGTGAAGAAAGAA 1220
DB 1738 TTGCACATGCTGCTGGCTACATCATCTCGGTTCCGATGATGAATCTGACGAGAGAG 1797
QY 1221 TTGTCAACCCCTCAAGAGTATAGCTCTTGTCTAATTGAG 1257
DB 1798 TTGTGTCCCTCAATCTATTTTGTCTAATTCTG 1834
```

RESULT 9
US-08-481-970-9
Sequence 9, Application US/08481970
Patent No. 5859310
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
APPLICANT: Salfield, Jochem
APPLICANT: Voss, Jeffrey
TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,970
FILING DATE:
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/260,452
FILING DATE: 14-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-013CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4963 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
IMMEDIATE SOURCE:
CLONE: PUHD BGR4
US-08-481-970-9

Query Match 4.6%; Score 136.6; DB 2; Length 4963;
Best Local Similarity 52.3%; Pred. No. 7,8e-27;
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

```
QY 507 TGTGTTAGTGTGTGTCATCGCTTCTGGTACCATATGSGGTAGCATCATGTGAAGC 566
DB 1018 TGTGCAAGTGTGAATGATGATGCTTCAAGCTACCATATGAGATCTGCTGTGAGGCG 1077
QY 567 TGCAGGCAATTTCTTAAGAGCAATTCAAGGCATATATGATATACAGTCCGCTCCACG 626
DB 1078 TGCAGGCTTCTTCAAGAGAGATATCAAGAGCATTAAGACATATATGATGTCCAGCCAC 1137
```

```

QY 627 AATGATGTGAATCAAAAGCGCAGACGTAAATCTGCGCAAGGCTTGCGCTTCATGAAG 686
Db 1138 AACCAAGTGCACCATGTATATAAAACAGAGAGAAAGACCTGCGCGCTCCGCCAAA 1197
QY 687 TGTTTAAAGTGGGATGTGTGAAGAAAGGGTGGCTCTTGACAGAGTACGTGAGAGCTGG 746
Db 1198 TGTCTAGAGTGGGATGTATGAAAGTGGGATACCAAAAAGCCGAAGAGGAGGAATG 1257
QY 747 CAGAAGTCAAGCGCAGATAGATGCGGAGAA----- 778
Db 1258 TTGAACACAAAGCCGCAGAGATATATGGGAGGCGCAGGGGTGAAGTGGGGTCTGCTGGA 1317
QY 779 -----CAGCCCATACCTGAACCTTCACTGCTGTGACGCCAGCAAAAA----- 820
Db 1318 GACATGAGAGCTGCCAACCTTTGGCCAAAGCCCGCTCATGATCAAAACGCTTAAGAAAGAC 1377
QY 821 -----GCCATATACAGAGTTGTCTCAATTTGTTGGTGGCTGGAACCG 863
Db 1378 AGCTTGCCCTTGCTCCCTGACCGCGACCAAGTGTCAATGAGCTGCTGAGGCC 1437
QY 864 GAGAAAGTCTATGCGCATGCTGACCTGACTGTGCTCCCGCAGTGAATCAAAAGCCCTCACT 923
Db 1438 CCGATCTCTATTTCCAGATATATCTTACAGACCTTCAAGTGAAGCTTCATGATGGGC 1497
QY 924 AACTGTGTGACTTGGCCGACCGAAGTGTGTGTTATCATTTGATGGGCGAAGCATATT 983
Db 1498 TTACTGACCAACCTGGCAGACAGGAGAGCTGTTCATCATGATCAACTGGGCGAAGAGGCTG 1557
QY 984 CCAGGCTTTCACAGCTGTGCTGCGGCGGACCAAGATGAGCTTCTGCAAGTCTTGATG 1043
Db 1558 CCAGGCTTGTGATTTGATTTGACCTTCATGATCAAGCTTCAAGATGAGTGGCTGAGCTA 1617
QY 1044 GAAATTTGATCTTGTGTGTGTATACCGGTCTCTTTCATTTGAGATGAGATGCTGTAT 1103
Db 1618 GAGATCTGTATGATTTGGTCTGTCTGCGCTCCATGAGACCCAGTGAAGCTACTGTTT 1677
QY 1104 GCAGACATTAATATATGAGCAAGAAAGCATGCAAAAT---TAGCAGGCTTCTTGATCTA 1160
Db 1678 GCTCTTAAGTGTCTTGGACAGAGAACAGGAAATGTGTAGAGGGATGTGTGAGATC 1737
QY 1161 AATAATGCTATCTGCGACGTGTGAAGAAATACAAAGCATGAGCTGGAAGAAAGAA 1220
Db 1738 TTGCACATGCTGCTGGCTACATCATCTCGTTCGCAATGATATCTGCAAGGAGAGAG 1797
QY 1221 TTGTCAACCTCAAGCATATAGCTCTTGCTAATTCAG 1257
Db 1798 TTGTGTGCTCAAAATCTATTTTGTCTAATTTCTG 1834

```

```

RESULT 10
US-08-897-719-9
; Sequence 9, Application US/08897719
; Patent No. 5923927
;
GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bayard, Hermann
; APPLICANT: Salfeld, Jochem
; APPLICANT: Voss, Jeffrey
; TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
; TITLE OF INVENTION: Cells by Tetraacycline-responsive Promoters
; NUMBER OF SEQUENCES: 10
;
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Iahive & Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,719
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/260,452
FILING DATE:
APPLICATION NUMBER: 08/076,327
FILING DATE: 14-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BHI-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4963 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Human cytomegalovirus
IMMEDIATE SOURCE:
CLONE: pUHD B64

```

US-08-897-719-9

Query Match 4.6%; Score 136.6; DB 2; Length 4963;

Best Local Similarity 52.3%; Pred. No. 7.8e-27;

Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

```

QY 507 TGTTTAGTGTGTGATGATGCTTCTGCGTACCACTATGAGGATGATGATGAGAGCC 566
Db 1018 TGTCAAGTGTGATGATGATGCTTCTGCGTACCACTATGAGGATGATGATGAGAGCC 1077
QY 567 TGCAGGCTTCTTCAAGAGAGCAATTCAGGCAATATGATATCAGTGTCTGCGCAG 626
Db 1078 TGCAGGCTTCTTCAAGAGAGCAATTCAGGCAATATGATATGATGATGCGCAGC 1137
QY 627 AATGATGTGAATCAAAAGCGCAGACGTAAATCTGCGGCTTGCGCTTCATGAAG 686
Db 1138 AACCAAGTGCACCATGTATATAAAACAGAGAGAAAGACCTGCGCGCTCCGCCAAA 1197
QY 687 TGTTTAAAGTGGGATGTGTGAAGAAAGGGTGGCTCTTGACAGAGTACGTGAGAGCTGG 746
Db 1198 TGTCTAGAGTGGGATGTATGAAAGTGGGATACCAAAAAGCCGAAGAGGAGGAATG 1257
QY 747 CAGAAGTCAAGCGCAGATAGATGCGGAGAA----- 778
Db 1258 TTGAACACAAAGCCGCAGAGATATATGGGAGGCGCAGGGGTGAAGTGGGGTCTGCTGGA 1317
QY 779 -----CAGCCCATACCTGAACCTTCACTGCTGTGACGCCAGCAAAAA----- 820
Db 1318 GACATGAGAGCTGCCAACCTTTGGCCAAAGCCCGCTCATGATCAAAAGCTTCAAGAAAGAC 1377
QY 821 -----GCCATATACAGAGTTGTCTCAATTTGTTGGTGGCTGGAACCG 863
Db 1378 AGCTTGCCCTTGCTCCCTGACCGCGACCAAGTGTCAATGAGCTGCTGAGGCC 1437
QY 864 GAGAAAGTCTATGCGCATGCTGACCTGACTGTGCTCCCGCAGTGAATCAAAAGCCCTCACT 923
Db 1438 CCGATCTCTATTTCCAGATATATCTTACAGACCTTCAAGTGAAGCTTCATGATGGGC 1497
QY 924 AACTGTGTGACTTGGCCGACCGAAGTGTGTGTTATCATTTGATGGGCGAAGCATATT 983
Db 1498 TTACTGACCAACCTGGGAGAGAGGAGCTGTTCATCATGATCAACTGGGCGAAGAGGCTG 1557
QY 984 CCAGGCTTTCACAGCTGTGCTGCGGACCAAGATGAGCTTCTGCAAGTCTTGATG 1043
Db 1558 CCAGGCTTGTGATTTGACCTTCAAGATGATGATGATGATGATGATGATGATGATGATG 1617

```


	Matches	427,	Conservative	0;	Mismatches	324,	Indels	66;	Gaps	3
OY	507	TGTTTAGTGTGGTGACATCGCTTTGGSGTAACAATAAGGGSTACCATTCAATGTTGAAGCC								566
Dd	1018	TGTGCAGTGTGCAAATGACTATGCTTCAGGGCTAACCATTAATGAGGTCTGGTCTGTGAGGGC								1077
OY	567	TGCAGAGCANNTCTTCAAAGAGA CAATTCAAAGCAATFAGAAATACAGSCTGCCCTGCACG								626
Dd	1078	TGCAGAGCCCTCTTCAAAGAAAGTATTCAAAGACATPAAGACATPATATGTGTCCAGCACCC								1137
OY	627	AATGAATGTGMMAATCAAAAAGCGCAGACGTAAATCTGTGCCAGGCTTTCCGCTTCATGAAG								686
Dd	1138	AACCACTGACCACTTATATMAAACAAGAGBABAAGCTGCCAGGCTCTCCGGCTCCGCAAA								1197
OY	687	TGTTTTAAAGTGGGAC TGCTGAAAAGAGGGGTGC GTTGTGACAGATACGTGANAGTTCG								746
Dd	1198	TGCTTACGAAGTGGGAATGATGAAAGGTGGGGATACGA AAAGCCGAAGAGAGGAGGAATG								1257
OY	747	CBAAGTACAAAGGCGAGATAGATGGGGAA-----								778
Dd	1258	TTGAAACA CAGAGCGCAGAGATGATGGGAGGAGCGAGSGGTGAAGTGGGCTCTGTGGA								1317
OY	779	-----CAGCCATACCTGMAACCTCA GCTGAGTTCA GCCAGCCA AAA- --								820
Dd	1318	GACATGAGAGCTGCCAACCTTTGGC AAGCCCGT CATGATCA AACGCTCTAAGAA GAAC								1377
OY	821	-----GCCATATACAAATTTGTCTCA ATTTGTTGGTGGCTGAACCG								863
Dd	1378	AGCCTGGCCTTGTCCCCTGAGGGCGCACAATAGTGTCA TGGCCTTGTGGATGTGTGAGCCC								1437
OY	864	GAGAAAGTCTATGCGCATNGCTGA ACCTACGTCCCCGACAGTGA CATCAAGCCCTCACT								923
Dd	1438	CCCATACTTATTCGAGATGATFTCTTACAGACACCTTCA GTGAAAGCTTGTGATGAGGC								1497
OY	924	ACA CTGTGTACTTTGSCCGACCCGAGAGTGTGTTATCAT TTGATGGGCGAGACATATT								983
Dd	1498	TTACTGACCAACCTCGGAGACAGGGA CTCGTTCACTGATCA ACTGGCGCAAGAGGGGT								1557
OY	984	CCAGGCTTTCACAGCTGTCCCTGGCGG ACAGATGAGCTTCTG CAGAGTCTTGANTG								1043
Dd	1558	CCAGGCTTTTGAGATTTGACCCCTTCATGATCAGGTTCCACCTTTCAGATGTGCCTGGCTA								1617
OY	1044	GAA TTTTGANOCCTGSHNCOGNAPACO GGCTCTTTCA TTTGAGATGA ACTGTCTAT								1103
Dd	1618	GAGATCCTGATGATTTGTGTCTGTGGGCTCCATGAGAC CCAGTGAAGCTACTGTIT								1677
OY	1104	GCAAGCATTTATATATGACGAGAACCA AGTCCAAAT-- TAGCAGGCTTCTTGATCTA								1160
Dd	1678	GCTCCTTAAGTCTGCTTGGAGAGGAA CCAAGGAAAAATGTGTAAGGGGCA TGTGSAATTC								1737
OY	1161	AATAATGCTATCTTCAGCTGTGTAAAGAAATACAGAGCATGAAGCTGGAAAAAGAAATA								1222
Dd	1738	TTGCACATGTGCTGGGTACATCATCTTCGGTTCCG CATGATGAATCTGCAGGAGAGGANG								1797
OY	1221	TTTGTCACCCCTCAAGAAGCTATGACTCTTGCTATTTACAG								1257
Dd	1798	TTTGtggctcCaAtCTATtATttTtgctTAAttcg								1834
 RESULT 12 US-09-041-886-34 ; Sequence 34, Application US/09041886 ; Patent No. 6235872 ; GENERAL INFORMATION: ; APPLICANT: Bredezen, Dale E. ; APPLICANT: Rabizadeh, Sharroz ; TITLE OF INVENTION: Proapoptotic Peptides, Dependence ; TITLE OF INVENTION: Polypeptides and Methods of Use ; NUMBER OF SEQUENCES: 72 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: Campbell & Flores LLP ; STREET: 4370 La Jolla Village Drive, Suite 700 ; CITY: San Diego										

STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-1J 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 6450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 361..2146
US-09-041-886-34

Query Match 4.6%; Score 136.6; DB 3; Length 6450;
Best Local Similarity 52.3%; Pred. No. 9.1e-27;
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

507 TGTGTTAGTGTGTGTCATCGCTTCTGGGTACACATATGAGATGTTGAAGCC 566
913 TGTGACGTGTGCAATGACTATGCTTCAAGGCTACCATTAAGAGTGTGAGAGGC 972
567 TGCAGGCAATCTTCAGAGGCAATTCAGGCAATATAGAAATCAGCTGCCCGCCAGC 626
973 TGCAGGCTTCTTCAGAGGCAATTCAGGCAATATAGAAATCAGCTGCCCGCCAGC 1032
627 AATGATGTGAATCAGAAAGCGAGAGCTGAATCTCCAGGCTTCCGCTTCATGAG 686
1033 AACGAGTGCACCAATTAATAAACAAGAGAGAGAGCTGCCAGGCTCCGCTCCGAAA 1092
687 TGTGTTAAAGTGGGCAATGCTGAAGAGAGGAGGCTGCTTACAGAGATACGTGAGTCGG 746
1093 TGTACGAAAGTGGGAAATGATGAAGAGGAGGATGCGAAGAACCGAAGAGAGAGATG 1152
747 CAGAGATCAAGGCGAGAGATAGTCGAGAA----- 778
1153 TTGAACAACAAGGCGCAGAGATAGTGGGAGGCGAGGAGTGAAGTGGGCTCTGCTGGA 1212
779 -----CAGCCATACCTGAACCTTCTAGCTGTTAGCCAGCAAAAA--- 820
1213 GACATGAGAGCTGCAACCTTTGCGCAAGCCCGCTCATGATCAAAAGCTCTAAGAGAAC 1272
821 -----GCCATATACAGATTTGTCTCAATTTGTGGTGGTGAACCG 863
1273 AGCTGGCTTGTCCCTGACGCGCGCAGACAGATGATGATGCTGTTGGATGCTGAGCCC 1332
864 GAGAGATCTATGCTGACCTGACCTGACCTTCTCCCGACAGTGAATCAAGAGCCCTCAT 923
1333 CCATATCTATTCGAGATATGATCTTCAAGACCTTCTGATAGAGCTTGATGATGAGGCG 1392
924 AACCTGTGTGACTTGGCCGACGAGAGTGTGTGTTATCATTTGATGAGGAGCATATT 983
1393 TTACTGACCAACCTGGCAGAGAGAGCTGTTCATCATGATCAACTGGGCGAAGAGGATG 1452
984 CCAGGCTTCCACAGCTGTCCTCGCGGACCAATAGGCTTTGTGAGAGTGTGATG 1043
|||||

Db 1453 CAGGCTTTGTGATTTGACCCCTCCATGATCAGGTCACCTTTGATGATGCTCGCTA 1512
Qy 1044 GAATTTGATCCTTGTGTGTGATACCGGTCTCTTTCATTTGAGATGAATTGTCTAT 1103
Db 1513 GAGATCCGATGATGCTGTCTGTCTGCGCTCCATGAGCACCAGATGAAGCTACTGTTT 1572
Qy 1104 GCAGAGATTTATTAATGAGAGAGAGCCAGTCCAAAT---TAGAGGCTCTTGATCTA 1160
Db 1573 GCTCTTAAGCTTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1632
Qy 1161 AATAATGCTATCTGACGCTGTGTAAGAAATATCAAGAGCATGAAGCTGGAAGAGAA 1220
Db 1633 TTGCATAGCTGTCTGTGATCATCTCCGCTTCGAGATGATGATCTCAGAGAGAGAG 1692
Qy 1221 TTTGTCACCTTCAAGATATGCTTGTCTATTTAG 1257
Db 1693 TTTGTGTGCTCAATCATATTATTTGCTTATTTCTG 1729

RESULT 13

US-08-453-998-1
Sequence 1, Application US/08453998
Patent No. 644438

GENERAL INFORMATION:
APPLICANT: CHAMON, PIERRE
APPLICANT: METZGER, DANIEL
APPLICANT: WHITE, JOHN
TITLE OF INVENTION: METHOD FOR THE PREPARATION OF A PROTEIN
TITLE OF INVENTION: BY YEASTS USING AN INDUCIBLE SYSTEM, VECTORS AND
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSER: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 NEW YORK AVE., N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,998
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/161,064
FILING DATE: 03-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: CHAPIN, MARIANA K.
REGISTRATION NUMBER: 35,843
REFERENCE/DOCKET NUMBER: 1037/98493
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3711
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6450 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-453-998-1

Query Match 4.6%; Score 136.6; DB 4; Length 6450;
Best Local Similarity 52.3%; Pred. No. 9.1e-27;
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;
507 TGTGTTAGTGTGTGTCATCGCTTCTGGGTACACATATGAGATGTTGAAGCC 566
|||||

```

Db      913  TGTGACGTGTGACATGCTATGCTTCAGGCTACCATTAATGAGTCTGCTCTGTGAGGC 972
Qy      567  TGCAGGCACTTCTTCAAGAGGACCAATTCAGGCAATATAGAAATACAGCTCCCTGCCACG 626
Db      973  TGCAGGCTCTTCTTCAAGAGGAGATTCAGAGCATTAACAGCTATATGTGTCCAGCCACC 1032
Qy      627  AATGAATGTGAATACAAAACGAGAGATTAATCTGCGAGGCTTCCGCTTCATGAG 686
Db      1033  AACCAAGTGCACCAATTGATTAATAACAGAGGAGAGCTGCCAGGCTCCGCTCCGCAAA 1092
Qy      687  TGTTTAAAAGTGGGCAATGCTGAAGAAGAGGCTGCTCTTGAACAAGTACAGAGGCTCG 746
Db      1093  TGTCTACAGAGTGGGAATGATGAAGGTGGATACGAAAAGCCGAAAGAGAGGAGAAATG 1152
Qy      747  CAGAAATGACAAAGCGCAGAGATGATGCGAGAA----- 778
Db      1153  TTGAAAACAGAGCCCGAGAGAGATGATGGGAGGCGAGGGGTGAAGTGGGCTGTGGA 1212
Qy      779  -----CAGCCATACCTGAACCTTGAAGCTGTGTTCAAGCCAGCAAAA----- 820
Db      1213  GACATGAGAGCTGCAACCTTTGGCCAAAGCCGCTCATGATCAAAAGCTTAAAGAGAAC 1272
Qy      821  -----GCCATATAACAAGATGTGTCTCAATTTGTTGGTGGCTGAACCG 863
Db      1273  AGCCTGGCTTGTCCCTGACGCGCAGACAGATGTGTGAGTGTGTGATGTGAGGCC 1332
Qy      864  GAGAAGATCTATGCGCATGCTGACCTTACCTGCTCCCGACAGTGAACATCAAAAGCCCTCACT 923
Db      1333  CCCTACTCTTATTCGAGATATGATCTTACAGACCTTCAAGTGAAGCTTCATGATAGGCG 1392
Qy      924  AACTGTGTGACTTGGCCGAGAGGTTGGTGTATTCATTTGATGGGCGAAGCATATT 983
Db      1393  TTACTGACCAACCTGGCAGACAGGAGCTGTTCACATGATCACTGGGCGAAGAGGGTGT 1452
Qy      984  CCAGGCTTCTCAGAGCTGTCCCTGGCGGACCAAGATGAAGCTTCTGCAAGAGCTTGGAGT 1043
Db      1453  CGAGCTTGTGGATTTAACCCTCATATACAGGTGACCTTCTGAAGATGTGCTGGCTA 1512
Qy      1044  GAAATTTTGTATCTTGTGTGTGTATACGGTCTCTTCAATTTGAGGATGAACCTTGTAT 1103
Db      1513  GAGATCTGATATGATGTGTCTGTCTGGGCTCCATGAGGACCCAGTGAAGCTTACTGTT 1572
Qy      1104  GCAGACGATTAATATATGAGACGAGACCAATTCGCAAGGCTTCTTATCTA 1160
Db      1573  GCTCTTAATCTGCTTGTGACAGAAACAGGAGAAATGTGAGAGGCAATGTGAGATC 1632
Qy      1161  AATATATGCTATCTGCAAGCTGTGAAGAAATCAAGAGCATGAAGCTGAAAAAGAGAA 1220
Db      1633  TTGACATGCTGTGCTACATCATCTCGGTTCCGATGATGAATCTGAGGAGAGGAG 1692
Qy      1221  TTTGTCAACCTCAAGCTATAGCTTTGCTAATTGAG 1257
Db      1693  TTTGTGTGCTCAAAATCTATTTTGTCTTAATTCG 1729

RESULT 14
US-10-027-983-3
; Sequence 3, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; APPLICANT: Mark P. Roach
; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
; FILE REFERENCE: RTS-0340
; CURRENT APPLICATION NUMBER: US/10/027,983
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 3
; LENGTH: 6450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

```

```

; LOCATION: (361) ... (2148)
US-10-027-983-3
Query Match      4.6%; Score 136.6; DB 4; Length 6450;
Best Local Similarity 52.3%; Pred. No. 9.1e-27;
Matches 427; Conservative 0; Mismatches 324; Indels 66; Gaps 3;

Qy      507  TGTATATGTGTGTGATGATCAGTCCTTCTGAGTACCACTATGGGGAGAGCATATGAAAGCC 566
Db      913  TGTGACGTGTGACATGCTATGCTTCAGGCTACCATTAATGAGTCTGCTCTGTGAGGC 972
Qy      627  AATGAATGTGAATACAAAACGAGAGATTAATCTGCGAGGCTTCCGCTTCATGAG 686
Db      1033  AACCAAGTGCACCAATTGATTAATAACAGAGGAGAGAGCTGCCAGGCTCCGCTCCGCAAA 1092
Qy      687  TGTTTAAAAGTGGGCAATGCTGAAGAAGAGGCTGCTCTTGAACAAGTACAGAGGCTCG 746
Db      973  TGCAGGCTCTTCTTCAAGAGAGATTCAGAGCATTAACAGCTATATGTGTCCAGCCACC 1032
Qy      747  CAGAAATGACAAAGCGCAGAGATGATGCGAGAA----- 778
Db      1153  TTGAAAACAGAGCCCGAGAGAGATGATGGGAGGCGAGGGGTGAAGTGGGCTGTGGA 1212
Qy      779  -----CAGCCATACCTGAACCTTGAAGCTGTGTTCAAGCCAGCAAAA----- 820
Db      1213  GACATGAGAGCTGCAACCTTTGGCCAAAGCCGCTCATGATCAAAAGCTTAAAGAGAAC 1272
Qy      821  -----GCCATATAACAAGATGTGTCTCAATTTGTTGGTGGCTGAACCG 863
Db      1273  AGCCTGGCTTGTCCCTGACGCGCAGACAGATGTGTGAGTGTGTGATGTGAGGCC 1332
Qy      864  GAGAAGATCTATGCGCATGCTGACCTTACCTGCTCCCGACAGTGAACATCAAAAGCCCTCACT 923
Db      1333  CCCTACTCTTATTCGAGATATGATCTTACAGACCTTCAAGTGAAGCTTCATGATAGGCG 1392
Qy      924  AACTGTGTGACTTGGCCGAGAGGTTGGTGTATTCATTTGATGGGCGAAGCATATT 983
Db      1393  TTACTGACCAACCTGGCAGACAGGAGCTGTTCACATGATCACTGGGCGAAGAGGGTGT 1452
Qy      984  CCAGGCTTCTCAGAGCTGTCCCTGGCGGACCAAGATGAAGCTTCTGCAAGAGCTTGGAGT 1043
Db      1453  CGAGCTTGTGGATTTAACCCTCATATACAGGTGACCTTCTGAAGATGTGCTGGCTA 1512
Qy      1044  GAAATTTTGTATCTTGTGTGTGTATACGGTCTCTTCAATTTGAGGATGAACCTTGTAT 1103
Db      1513  GAGATCTGATATGATGTGTCTGTCTGGGCTCCATGAGGACCCAGTGAAGCTTACTGTT 1572
Qy      1104  GCAGACGATTAATATATGAGACGAGACCAATTCGCAAGGCTTCTTATCTA 1160
Db      1573  GCTCTTAATCTGCTTGTGACAGAAACAGGAGAAATGTGAGAGGCAATGTGAGATC 1632
Qy      1161  AATATATGCTATCTGCAAGCTGTGAAGAAATCAAGAGCATGAAGCTGAAAAAGAGAA 1220
Db      1633  TTGACATGCTGTGCTACATCATCTCGGTTCCGATGATGAATCTGAGGAGAGGAG 1692
Qy      1221  TTTGTCAACCTCAAGCTATAGCTTTGCTAATTGAG 1257
Db      1693  TTTGTGTGCTCAAAATCTATTTTGTCTTAATTCG 1729

RESULT 15
US-09-833-381-269/c
; Sequence 269, Application US/0983381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381

```

```
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 269
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-269

Query Match      4.4%; Score 131.6; DB 4; Length 472;
Best Local Similarity 58.3%; Pred. No. 4.6e-26;
Matches 275; Conservative 0; Mismatches 184; Indels 13; Gaps 2;

QY 1007 GGCAGACCCAGATGAGCCTTCTGCAAGATGCTGATGAGAAATTTGATCCTTGGTGTCT 1066
Db 472 GTCTGACCAAGATGCTAGTACTGCAAGAGCGTGTGATGAGAGTGTGCTGGTGGTGGTGC 413
QY 1067 ATACCGGCTCTTTTCATTGAGATGAACTTGTCTATGCAAGCATTTATATATGACGA 1126
Db 412 TCAGCGGCTCCCTGACGCTGACAGATGAGCTGGCTTGGCTGAGTACTTAGTCTGGATGA 353
QY 1127 AGACCAATCCAAATTAGAGGCGCTTCTGATCTAATA-ATGCTATCTGCACTGGTAA 1185
Db 352 AGAGGGGGCAGCGCCAGCTGAGATGGGGAAATGGGGAGCTGCCCTGCTGCACTAGTGC 293
QY 1186 AGAAATACAGAGCATGAGCTGAGAAAAAGAAATTTGTCAACCTCAAGCTATAGCTC 1245
Db 292 GTCGGCTGACAGGCGCTGCGATGAGCGATGAGATATGTTCTAGTAAAGGCTTGGGCC 233
QY 1246 TTGCTAATTCAGATCTCATGCACTAGAAAGATGTTGAAGCCGTTGAGAGCTTCAGGATG 1305
Db 232 TTGCAATTCAGATCTCATGCACTAGAAAGATGTTGATGTTGAGAGAGCTGCGAGAG 173
QY 1306 TCTTACATGAGAGCGCTGACAGATTATGAGCTGGCCAG-----CACTGGAAG 1353
Db 172 CTATGCAAGAGGCGCTGATGATGATATCCGAGGCGCTGGCACCGGAGGGGTTTG 113
QY 1354 ACCCTGCTGAGCTGCGAAGATGCTGATGACACTGCCACTCCTGAGGCAAGCTTACCA 1413
Db 112 AACGTGGCGGGCGGGCAGGGTGTCTCAAGCTACCGCTCTCTCCGCAAGACCGGGCA 53
QY 1414 AGGCCGTCGACGATTTTCAACAATCAACTAGAAAGGCAAAAGTCCCAATGCA 1465
Db 52 AAGTGTCTGGCCCATTTCTATGAGGTGAAGCTGAGGGCAAGGTGCCCATGCA 1
```

Search completed: August 10, 2004, 04:59:14
Job time : 226 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2004, 23:22:43 ; Search time 2392 Seconds
(without alignments)
6118.676 Million cell updates/sec

Title: US-10-054-841-3
Perfect score: 2985
Sequence: 1 ggcggccgcagtggtggtg.....actggcgccgcgcgttgat 2985

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3222919 seqs, 245157024 residues
Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/ECT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2985	100.0	2985	US-10-054-841-3	Sequence 3, Appli
2	2973	99.6	2987	US-10-054-841-5	Sequence 5, Appli
3	1377	46.1	1377	US-10-355-218-3	Sequence 3, Appli
4	715	24.0	2807	US-10-054-841-1	Sequence 1, Appli
5	695	23.3	2153	US-10-101-510-87	Sequence 87, Appli
6	635	23.3	2153	US-10-101-510-413	Sequence 413, Appli
7	428	14.3	622	US-10-054-841-8	Sequence 8, Appli
8	424.4	14.2	2218	US-09-962-436-268	Sequence 268, Appli
9	424.4	14.2	2218	US-10-776-827-4	Sequence 4, Appli
10	421.4	14.1	2539	US-10-037-270-454	Sequence 454, App
11	421.4	14.1	2539	US-10-117-722-454	Sequence 454, App
12	421.4	14.1	4250	US-10-297-653-2	Sequence 2, Appli
13	401.6	13.5	2402	US-10-342-887-1041	Sequence 1041, Ap
14	401.6	13.5	2402	US-10-172-118-1041	Sequence 1041, Ap

15	401.6	13.5	2402	15	US-10-211-239-7	Sequence 7, Appli
16	401.6	13.5	2402	15	US-10-101-510-66	Sequence 66, Appli
17	360.4	12.1	1599	16	US-10-093-463-201	Sequence 201, App
C 18	204.6	6.9	403	15	US-10-054-841-7	Sequence 7, Appli
C 19	191	6.4	565	13	US-10-027-632-285776	Sequence 285776,
C 20	191	6.4	565	16	US-10-027-632-285776	Sequence 285776,
C 21	189.8	6.4	565	13	US-10-027-632-285777	Sequence 285777,
C 22	189.8	6.4	565	16	US-10-027-632-285777	Sequence 285777,
C 23	141.4	4.7	1978	15	US-10-052-092-23	Sequence 23, Appli
24	141.4	4.7	1978	15	US-10-437-107-23	Sequence 23, Appli
25	136.6	4.6	1785	15	US-10-095-373A-1	Sequence 1, Appli
26	136.6	4.6	1788	9	US-09-853-033-1	Sequence 1, Appli
27	136.6	4.6	2092	10	US-09-952-680A-9	Sequence 9, Appli
28	136.6	4.6	2092	15	US-10-052-092-2	Sequence 6, Appli
29	136.6	4.6	2092	15	US-10-157-899A-54	Sequence 54, Appli
30	136.6	4.6	2092	15	US-10-437-107-6	Sequence 6, Appli
31	136.6	4.6	2106	15	US-10-052-092-5	Sequence 5, Appli
32	136.6	4.6	2106	15	US-10-437-107-5	Sequence 5, Appli
33	136.6	4.6	4963	9	US-09-281-674-9	Sequence 9, Appli
34	136.6	4.6	4963	9	US-09-892-227-9	Sequence 9, Appli
35	136.6	4.6	4963	13	US-09-777-317-9	Sequence 43, App
36	136.6	4.6	6450	13	US-10-342-887-435	Sequence 435, App
37	136.6	4.6	6450	13	US-10-172-118-435	Sequence 403, App
38	136.6	4.6	6450	13	US-10-388-360-303	Sequence 3, Appli
39	136.6	4.6	6450	13	US-10-448-753-3	Sequence 2, Appli
40	136.6	4.6	6450	14	US-10-096-710-2	Sequence 1, Appli
41	136.6	4.6	6450	14	US-10-081-563-1	Sequence 1, Appli
42	136.6	4.6	6450	15	US-10-052-092-1	Sequence 1, Appli
43	136.6	4.6	6450	15	US-10-052-092-7	Sequence 7, Appli
44	136.6	4.6	6450	15	US-10-207-655-60	Sequence 60, Appli
45	136.6	4.6	6450	15	US-10-177-233-127	Sequence 127, App

ALIGNMENTS

RESULT 1
US-10-054-841-3
; Sequence 3, Application US/10054841
; Publication No. US20030119100A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: 19999YCA
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: 09/487,379
; PRIOR FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/141,000
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/078,633
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/062,902
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/057,090
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2985
; TYPE: DNA
; ORGANISM: Human
US-10-054-841-3

Query Match 100.0%; Score 2985; DB 15; Length 2985;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGGCGCCAGTGTGTGATTCGCTTGCATCAGAGCAACATTTGTTAATTGCA 60
Db 1 GCGGGCGCCAGTGTGTGATTCGCTTGCATCAGAGCAACATTTGTTAATTGCA 60

QY 61 CTGTGCTCTGTCAGAGAACTTTGATTATAGTGGGGTGCAAAATATGTTGCCGT 120
DB 61 CTGTGCTCTGTCAGAGAACTTTGATTATAGTGGGGTGCAAAATATGTTGCCGT 120
QY 121 CGCATATGATTTGGGTAGAACTTTGCTCTGAACTTTTTCCTGCACTACGAGAAAG 180
DB 121 CGCATATGATTTGGGTAGAACTTTGCTCTGAACTTTTTCCTGCACTACGAGAAAG 180
QY 181 AGCTTCTGTGAGAAATGTCAACAAAGATGACACATTTGCTGAGCTGTGCTCTTCA 240
DB 181 AGCTTCTGTGAGAAATGTCAACAAAGATGACACATTTGCTGAGCTGTGCTCTTCA 240
QY 241 TCAGAGCGAACTTTGCAAGCCGCTGCTGACGAGCAGGTCACACCAAGCCCTG 300
DB 241 TCAGAGCGAACTTTGCAAGCCGCTGCTGACGAGCAGGTCACACCAAGCCCTG 300
QY 301 GTGGCTCTTTCAGAGCCGCTGAGGAGCTACGTTCAACATATGAGCCATCAGAACGAG 360
DB 301 GTGGCTCTTTCAGAGCCGCTGAGGAGCTACGTTCAACATATGAGCCATCAGAACGAG 360
QY 361 TTGACTGCGCACTCTCTACCTCTGCTGCTGCTGCTGAGGATGAGGGCTGTGAGGA 420
DB 361 TTGACTGCGCACTCTCTACCTCTGCTGCTGCTGCTGAGGATGAGGGCTGTGAGGA 420
QY 421 AACTGTATGATGACTGCTGCGACCAATTTGTAAGATCCCAAGCTGTAATACA 480
DB 421 AACTGTATGATGACTGCTGCGACCAATTTGTAAGATCCCAAGCTGTAATACA 480
QY 481 TGCTCACTGATGAGCCCAAGAGACTGTGTAGGTGTGAGACATGCTTCTGAGTACC 540
DB 481 TGCTCACTGATGAGCCCAAGAGACTGTGTAGGTGTGAGACATGCTTCTGAGTACC 540
QY 541 ACTATGGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 ACTATGGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 ATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 CTTGCGAGGCTTGGCGCTTCTATGAAAGTGTAAAGTGTGAGATGCTGTAAGAGGGTGC 720
DB 661 CTTGCGAGGCTTGGCGCTTCTATGAAAGTGTAAAGTGTGAGATGCTGTAAGAGGGTGC 720
QY 721 GTCTTGAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 GTCTTGAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCCCATACCTGAACTCTGAGCTGATGAGCCGCAAAAGCCATATACAGATTTGCT 840
DB 781 GCCCATACCTGAACTCTGAGCTGATGAGCCGCAAAAGCCATATACAGATTTGCT 840
QY 841 CACATTTGTTGTTGGTGGAGACCGGAGAAAGATTTGTCATGCTGACCTTACTGCTCCCG 900
DB 841 CACATTTGTTGTTGGTGGAGACCGGAGAAAGATTTGTCATGCTGACCTTACTGCTCCCG 900
QY 901 ACAATGATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACGAGAGTGTGTTA 960
DB 901 ACAATGATCAAAAGCCCTCACTACACTGTGTGACTTGGCCGACGAGAGTGTGTTA 960
QY 961 TCATTTGATGAGGAGAGATATTCAGAGCTTCTCAAGCTGTCCCTGGCGGACAGATGA 1020
DB 961 TCATTTGATGAGGAGAGATATTCAGAGCTTCTCAAGCTGTCCCTGGCGGACAGATGA 1020
QY 1021 GCCCTTCTGAGAGTGTGATGAGAAATTTTGAATCCCTTGTGTGCTATACCGGCTCTTT 1080
DB 1021 GCCCTTCTGAGAGTGTGATGAGAAATTTTGAATCCCTTGTGTGCTATACCGGCTCTTT 1080
QY 1081 CATTTGAGATGAACTGTCTATGACAGAGATTAATATATATGAGAGCAAGCTCAAT 1140
DB 1081 CATTTGAGATGAACTGTCTATGACAGAGATTAATATATATGAGAGCAAGCTCAAT 1140
QY 1141 TAGCAGGCTTCTTGTATCTTAATATATGCTATCCTGACGCTGTAAAGAAATACAGAGCA 1200
DB 1141 TAGCAGGCTTCTTGTATCTTAATATATGCTATCCTGACGCTGTAAAGAAATACAGAGCA 1200
QY 1201 TGAAGCTGAAAAAGAAATTTTGTACCTTCAAGCTATATAGCTTTTCTATTCAGACT 1260
DB 1201 TGAAGCTGAAAAAGAAATTTTGTACCTTCAAGCTATATAGCTTTTCTATTCAGACT 1260
QY 1261 CCATGCAATATGAAAGATTTGAAAGCCGTTCAGAAAGCTTCAAGATGTCTTACATGAGAGC 1320
DB 1261 CCATGCAATATGAAAGATTTGAAAGCCGTTCAGAAAGCTTCAAGATGTCTTACATGAGAGC 1320
QY 1321 TGCAAGATTAATGAAAGCTGAGCAGCATGAGAAAGCCCTGCTGAGCTGGCAAGATGCTGA 1380
DB 1321 TGCAAGATTAATGAAAGCTGAGCAGCATGAGAAAGCCCTGCTGAGCTGGCAAGATGCTGA 1380
QY 1381 TGAACCTGCACTCTGAGGAGACCTTCTACCAAGGCGGTGACATTTCTTACAAATCA 1440
DB 1381 TGAACCTGCACTCTGAGGAGACCTTCTACCAAGGCGGTGACATTTCTTACAAATCA 1440
QY 1441 AACTATGAGGAGAAAGTCCCATGCAAAACCTTTTGGAAATGTGAGAGCCAAAGCT 1500
DB 1441 AACTATGAGGAGAAAGTCCCATGCAAAACCTTTTGGAAATGTGAGAGCCAAAGCT 1500
QY 1501 GACTTAAAGCTCCCTGGGCTTCCCATGCTTCACTGTTGAAAGAGGAGAAATPAAACCAAG 1560
DB 1501 GACTTAAAGCTCCCTGGGCTTCCCATGCTTCACTGTTGAAAGAGGAGAAATPAAACCAAG 1560
QY 1561 AGTATGTCGAGAAACTTATGAGATTTATGAAACATCAAAATCAACAGACTGACTG 1620
DB 1561 AGTATGTCGAGAAACTTATGAGATTTATGAAACATCAAAATCAACAGACTGACTG 1620
QY 1621 ATATTTATGAGAGAACTATGAGAGCTTTCAGATTTCTCCATAGCTTCTGATGAGT 1680
DB 1621 ATATTTATGAGAGAACTATGAGAGCTTTCAGATTTCTCCATAGCTTCTGATGAGT 1680
QY 1681 TCTTTACTTTTCCATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1740
DB 1681 TCTTTACTTTTCCATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1740
QY 1741 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1800
DB 1741 TTTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1800
QY 1801 CATTTCTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
DB 1801 CATTTCTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
QY 1861 TCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1920
DB 1861 TCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1920
QY 1921 AAAAATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
DB 1921 AAAAATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1980
QY 1981 TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2040
DB 1981 TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2040
QY 2041 CTTTAAAGAGGCTCTTAACTGAAAGAGATGAAAGCCAGCCCTGCAAGAGATGAGA 2100
DB 2041 CTTTAAAGAGGCTCTTAACTGAAAGAGATGAAAGCCAGCCCTGCAAGAGATGAGA 2100
QY 2101 TCCATATATGATGAGCAGGAACTTATGTGAACATACCTGCCCAATGACTTAAAGAA 2160
DB 2101 TCCATATATGATGAGCAGGAACTTATGTGAACATACCTGCCCAATGACTTAAAGAA 2160
QY 2161 TCAAAGAGAGAGAAACCAAGCTTCTTAAAGATGAGAGCAATATGAGAGT 2220
DB 2161 TCAAAGAGAGAGAAACCAAGCTTCTTAAAGATGAGAGCAATATGAGAGT 2220
QY 2221 GCAATTTGATTTCAATGAGAGAGCTCTTAAATTAAGCAACTTAAAGAGCTTGAATCGG 2280
DB 2221 GCAATTTGATTTCAATGAGAGAGCTCTTAAATTAAGCAACTTAAAGAGCTTGAATCGG 2280

Db 2221 GCAGTATTAGATTTCATGAGGACGACGCTCTAATTAGACAATTAGACAGCTTGCAATCGG 2280
Qy 2281 CTGCTTCCTTATCATTCAGTCTTTCATCTAGATCAGTTACAGCCATTGATTCCTTAATTTGT 2340
Db 2281 CTGCTTCCTTATCATTCAGTCTTTCATCTAGATCAGTTACAGCCATTGATTCCTTAATTTGT 2340
Qy 2341 TTTTTCAGTCTTCAGGTAATTTGTTAGTTTACTACTATGTAACTTTTTCAGGGAATAG 2400
Db 2341 TTTTTCAGTCTTCAGGTAATTTGTTAGTTTACTACTATGTAACTTTTTCAGGGAATAG 2400
Qy 2401 TTTTTCAGTCTTCAGGTAATTTGTTAGTTTACTACTATGTAACTTTTTCAGGGAATAG 2460
Db 2401 TTTTTCAGTCTTCAGGTAATTTGTTAGTTTACTACTATGTAACTTTTTCAGGGAATAG 2460
Qy 2461 GGCCTTGAACAATTAACGACATATAAGAGCAAAATGATCTGGAAGAGATTTTAA 2520
Db 2461 GGCCTTGAACAATTAACGACATATAAGAGCAAAATGATCTGGAAGAGATTTTAA 2520
Qy 2521 AAATGTTTGTCTTCTTCAAAATGAGATTTTGTGACCAAGCTTTTACCACTTTTCAG 2580
Db 2521 AAATGTTTGTCTTCTTCAAAATGAGATTTTGTGACCAAGCTTTTACCACTTTTCAG 2580
Qy 2581 CCATTTATTAATATGAGGAAATTTACTTACTCAAGCAATAGTTGAAGGAGAGTGATTT 2640
Db 2581 CCATTTATTAATATGAGGAAATTTACTTACTCAAGCAATAGTTGAAGGAGAGTGATTT 2640
Qy 2641 ATCCAGGATGCAATTTATGTTGTGTCAGGAGTCTGTCGCAAAATCATTTCTTAACTAG 2700
Db 2641 ATCCAGGATGCAATTTATGTTGTGTCAGGAGTCTGTCGCAAAATCATTTCTTAACTAG 2700
Qy 2701 AGCTCCAGTTTACCTTAATGTTGCTACTGACACAAAGATGAGTTTACCTTCACTGCTC 2760
Db 2701 AGCTCCAGTTTACCTTAATGTTGCTACTGACACAAAGATGAGTTTACCTTCACTGCTC 2760
Qy 2761 TGAGTAGTCACTATATATTAAGCACTGACATGAGATTTGATCCGTAAGATTTGTCAGGAGT 2820
Db 2761 TGAGTAGTCACTATATATTAAGCACTGACATGAGATTTGATCCGTAAGATTTGTCAGGAGT 2820
Qy 2821 GCACCTCTCTACTTGGAGAGTCAATTTGCTATGATTTCTAGCTGCAATGTTGTTAGG 2880
Db 2821 GCACCTCTCTACTTGGAGAGTCAATTTGCTATGATTTCTAGCTGCAATGTTGTTAGG 2880
Qy 2881 AATGTGATACCTGCTGTTTGGAAAGTCAAGACCTTTCCTGAGAAAGAGCTGTGAGCCAG 2940
Db 2881 AATGTGATACCTGCTGTTTGGAAAGTCAAGACCTTTCCTGAGAAAGAGCTGTGAGCCAG 2940
Qy 2941 TATTCATTTAAGAGAAATTCACCACTGCGCGCGCTTGAT 2985
Db 2941 TATTCATTTAAGAGAAATTCACCACTGCGCGCGCTTGAT 2985

RESULT 2
US-10-054-841-5
; Sequence 5; Application US/10054841
; Publication No. US20030119100A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: 19999YCA
; CURRENT APPLICATION NUMBER: US/10/054, 841
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: 09/487, 379
; PRIOR FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/141, 000
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/078, 633
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/062, 902
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/057, 090
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2987
; TYPE: DNA
; ORGANISM: Human
US-10-054-841-5
Query Match 99.6%; Score 2973; DB 15; Length 2987;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2985; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
Qy 1 GCGGCGCGCGAGTGTGAGAAATTCGCTTGTCACTAGGAAACATTTGTGTAATTGCA 60
Db 1 GCGGCGCGCGAGTGTGAGAAATTCGCTTGTCACTAGGAAACATTTGTGTAATTGCA 60
Qy 61 CTGTGCTCTGTCAAGAAATCTTGAATTTATAGCTGGGTGACAAATATATATGTTCCCGT 120
Db 61 CTGTGCTCTGTCAAGAAATCTTGAATTTATAGCTGGGTGACAAATATATGTTCCCGT 120
Qy 121 CGCAGATGATTCGGTAGAACTTTGCTTCTGAAATCTTTTCCCTGCACTACAGAGAG 180
Db 121 CGCAGATGATTCGGTAGAACTTTGCTTCTGAAATCTTTTCCCTGCACTACAGAGAG 180
Qy 181 AGCTTCTGTGAGAAATGCAAAAGATGCAACATTTGATTCAGCTGTTGCTCTTCA 240
Db 181 AGCTTCTGTGAGAAATGCAAAAGATGCAACATTTGATTCAGCTGTTGCTCTTCA 240
Qy 241 TCAAGACGGAACCTTCCAGCCAGCCCTCTGACGAGAGCGGTCAACCAAGCCCTG 300
Db 241 TCAAGACGGAACCTTCCAGCCAGCCCTCTGACGAGAGCGGTCAACCAAGCCCTG 300
Qy 301 GTGGCTCTTCAAGCGCCAGTGGAGTCAAGTTCAACATATGAGTGGATCGAAGC 360
Db 301 GTGGCTCTTCAAGCGCCAGTGGAGTCAAGTTCAACATATGAGTGGATCGAAGC 360
Qy 361 GTGGCTCTTCAAGCGCCAGTGGAGTCAAGTTCAACATATGAGTGGATCGAAGC 420
Db 361 GTGGCTCTTCAAGCGCCAGTGGAGTCAAGTTCAACATATGAGTGGATCGAAGC 420
Qy 421 AACTGTATGATGACTGCTCCAGACCAATTTGTAAGATTCGCCAGCAAGTGTGATACA 480
Db 421 AACTGTATGATGACTGCTCCAGACCAATTTGTAAGATTCGCCAGCAAGTGTGATACA 480
Qy 481 TGTCTAATCTGATGCTCCAGACCAATTTGTAAGATTCGCCAGCAAGTGTGATACA 540
Db 481 TGTCTAATCTGATGCTCCAGACCAATTTGTAAGATTCGCCAGCAAGTGTGATACA 540
Qy 541 ACTATGGGGTACATCATATGTAAGCTGCAAGGCAATTTCTTAAGAGACAAATTCAGGCA 600
Db 541 ACTATGGGGTACATCATATGTAAGCTGCAAGGCAATTTCTTAAGAGACAAATTCAGGCA 600
Qy 601 ATATGAATACAGCTGCCCTGCGACGAATGTAATTCACAAAGCGCAGACGTAAAT 660
Db 601 ATATGAATACAGCTGCCCTGCGACGAATGTAATTCACAAAGCGCAGACGTAAAT 660
Qy 661 CCTGCGAGCTTCCGCTTCATGAAAGTGTAAAGTGGGCATGCTGAAAGAAAGGGGTGC 720
Db 661 CCTGCGAGCTTCCGCTTCATGAAAGTGTAAAGTGGGCATGCTGAAAGAAAGGGGTGC 720
Qy 721 GTCTTGACAGAGTACGTGAGGTCGACAGAGTCAAGGCAAGGCAAGATGAGTGCAGAGACA 780
Db 721 GTCTTGACAGAGTACGTGAGGTCGACAGAGTCAAGGCAAGGCAAGATGAGTGCAGAGACA 780
Qy 781 GCCATACCTGAAACCTTCACTGCTTCAAGCCAGCCAAAAGCCATATTAACAAGATTTCT 840
Db 781 GCCATACCTGAAACCTTCACTGCTTCAAGCCAGCCAAAAGCCATATTAACAAGATTTCT 840
Qy 841 CATATTTGTTGGTGTGAACCGGAGAAATCTATGCTGCTGACCTTCACTGCTCCCG 900
Db 841 CATATTTGTTGGTGTGAACCGGAGAAATCTATGCTGCTGACCTTCACTGCTCCCG 900
Qy 901 ACAGTGACATCAAAAGCCCTCACTACCTGTGATGCTGCGCAGAGATTTGGTGTGA 960
Db 901 ACAGTGACATCAAAAGCCCTCACTACCTGTGATGCTGCGCAGAGATTTGGTGTGA 960

Db 901 AAGTGAATCAAAAGCCCTCACTACCTGTGACTTGGCGAGCCGAGGTTGTGTGTTA 960
 QY 961 TCATTGATGGAGGAGCATATTCAGGCTTCGCCAGCTGTCCCTGGCGAGCCAGATGA 1020
 Db 961 TCATTGATGGAGGAGCATATTCAGGCTTCGCCAGCTGTCCCTGGCGAGCCAGATGA 1020
 QY 1021 GCCTTTCGACAGTCTTGATGGAATTTTGTATCTTGTGTGTATACCGGTCTCTTT 1080
 Db 1021 GCCTTTCGACAGTCTTGATGGAATTTTGTATCTTGTGTGTATACCGGTCTCTTT 1080
 QY 1081 CATTTGGAGATGAGCTTGTCTATGACAGCATATATATATGAGAGACAGCTGCAAT 1140
 Db 1081 CATTTGGAGATGAGCTTGTCTATGACAGCATATATATATGAGAGACAGCTGCAAT 1140
 QY 1141 TAGAGGCTTCTTGATCTAATAATATGCTATCTGACAGCTGTAAAGAAATACAGAGCA 1200
 Db 1141 TAGAGGCTTCTTGATCTAATAATATGCTATCTGACAGCTGTAAAGAAATACAGAGCA 1200
 QY 1201 TGAAGCTGAAAAAGAGATTTGTACCCCTCAAGCTATAGCTTGTCTAATTCAGACT 1260
 Db 1201 TGAAGCTGAAAAAGAGATTTGTACCCCTCAAGCTATAGCTTGTCTAATTCAGACT 1260
 QY 1261 CCATGCAATGAGATGCTTGAAGCCGCTTCAGAGCTTCAGAGCTTCAGATGTCTTAATGAAGCC 1320
 Db 1261 CCATGCAATGAGATGCTTGAAGCCGCTTCAGAGCTTCAGAGCTTCAGATGTCTTAATGAAGCC 1320
 QY 1321 TGCAGGATTTATGAGCTGCGCAGCAGCATG - GAAGACCCCTGTGAGCTGSCAGAGTGTCT 1378
 Db 1321 TGCAGGATTTATGAGCTGCGCAGCAGCATGAGAGAGCCCTGTGAGCTGSCAGAGTGTCT 1378
 QY 1379 GATGACACTGCACTCTCTGAGGCGAGCCTCTACCAAGCCGCTGACAGCATTTCTAACAT 1438
 Db 1381 GATGACACTGCACTCTCTGAGGCGAGCCTCTACCAAGCCGCTGACAGCATTTCTAACAT 1440
 QY 1439 CAATCTGAGAGGCAAGTCCCAATGCACAAACTTTTGTGAATGTTGAGGCGCAAGT 1498
 Db 1441 CAATCTGAGAGGCAAGTCCCAATGCACAAACTTTTGTGAATGTTGAGGCGCAAGT 1500
 QY 1499 CTGACTAAAAGCTCCCTGGGCTTCCCATCTTCATGTTGAAAAAGGAAAAATTAACCA 1558
 Db 1501 CTGACTAAAAGCTCCCTGGGCTTCCCATCTTCATGTTGAAAAAGGAAAAATTAACCA 1560
 QY 1559 AGAGTGTATGCGAAGAACTTAGAGTTAGTTAACCAATCAAAATCAACAGACTGCAC 1618
 Db 1561 AGAGTGTATGCGAAGAACTTAGAGTTAGTTAACCAATCAAAATCAACAGACTGCAC 1620
 QY 1619 TGAATATTTAGAGAGAGCATATGAGAGCTTCAATTCCTCCATAGTCTCTATGA 1678
 Db 1621 TGAATATTTAGAGAGAGCATATGAGAGCTTCAATTCCTCCATAGTCTCTATGA 1680
 QY 1679 GTTCTTTCTAATCTTCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1738
 Db 1681 GTTCTTTCTAATCTTCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
 QY 1739 TATTTTTCCT 1798
 Db 1741 TATTTTTCCT 1800
 QY 1799 CCCATTCCT 1858
 Db 1801 CCCATTCCT 1860
 QY 1859 ATTCT 1918
 Db 1861 ATTCT 1920
 QY 1919 AAAAAAATCT 1978
 Db 1921 AAAAAAATCT 1980
 QY 1979 CT 2038
 Db 1981 CT 2040

QY 2039 AACTTTTAAAGAGTCTTCTAATGAGAGATGGAAGCCAGCCCTGCAAGAGATGA 2098
 Db 2041 AACTTTTAAAGAGTCTTCTAATGAGAGATGGAAGCCAGCCCTGCAAGAGATGA 2100
 QY 2099 GATCCATATATGATGAGCAGTGAACCTTATTTGTGAACCATACCCGCCCAATGACTAAGG 2158
 Db 2101 GATCCATATATGATGAGCAGTGAACCTTATTTGTGAACCATACCCGCCCAATGACTAAGG 2160
 QY 2159 AATCAAGAGAGAGAACCAAGTCTTAAAGTACATGCAACATATACAAATGACTGA 2218
 Db 2161 AATCAAGAGAGAGAACCAAGTCTTAAAGTACATGCAACATATACAAATGACTGA 2220
 QY 2219 GTGAGATTTAGATTTCAATGGAGAGCCCTTATTTAACAACCTTAAAGACCTTGACT 2278
 Db 2221 GTGAGATTTAGATTTCAATGGAGAGCCCTTATTTAACAACCTTAAAGACCTTGACT 2280
 QY 2279 GGCCTCTCTTATCATGCTTCTTCATCTAGATGATGATGATGATGATGATGATGATGAT 2338
 Db 2281 GGCCTCTCTTATCATGCTTCTTCATCTAGATGATGATGATGATGATGATGATGATGAT 2340
 QY 2339 GTTTTTCAAGTCTTCAAGGATTTTGTAGTTAGCTACTATGTATCTTTTTCAGGAGAT 2398
 Db 2341 GTTTTTCAAGTCTTCAAGGATTTTGTAGTTAGCTACTATGTATCTTTTTCAGGAGAT 2400
 QY 2399 AGTTAAGCTTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2458
 Db 2401 AGTTAAGCTTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
 QY 2459 CTGGCTTTGAACAATTAAGAACTAATGAGAGCAAAATGATCTGAGAGAGATTTT 2518
 Db 2461 CTGGCTTTGAACAATTAAGAACTAATGAGAGCAAAATGATCTGAGAGAGATTTT 2520
 QY 2519 AAAAAATGTTTGTCTTCTTCAAAATGAGATTTTGTGACCAAGCTTTTCAACTTTTC 2578
 Db 2521 AAAAAATGTTTGTCTTCTTCAAAATGAGATTTTGTGACCAAGCTTTTCAACTTTTC 2580
 QY 2579 AGCATTTATTAATGAGAAATTTAATCTTCAAGAAATGAGAGAGAGAGAGATGATA 2638
 Db 2581 AGCATTTATTAATGAGAAATTTAATCTTCAAGAAATGAGAGAGAGAGAGATGATA 2640
 QY 2639 TTAATCAAGATGAGAAATTTATGTTGTGAGAGTGTGAGAGAGAGAGAGAGATTTT 2698
 Db 2641 TTAATCAAGATGAGAAATTTATGTTGTGAGAGTGTGAGAGAGAGAGAGATTTT 2700
 QY 2699 TGAAGCTCAAGTTTAACTTAAATGTTCACTGACACAAAGATGAGATTAACCTTACAGTAC 2758
 Db 2701 TGAAGCTCAAGTTTAACTTAAATGTTCACTGACACAAAGATGAGATTAACCTTACAGTAC 2760
 QY 2759 TCTGAGATGATCAATATATATGAGAGCTGACATGATGATGATGATGATGATGATGATGAT 2818
 Db 2761 TCTGAGATGATCAATATATATGAGAGCTGACATGATGATGATGATGATGATGATGATGAT 2820
 QY 2819 GTGACCTCTCTAATGAGAGGATGAATGACATATGATGATGATGATGATGATGATGATGAT 2878
 Db 2821 GTGACCTCTCTAATGAGAGGATGAATGACATATGATGATGATGATGATGATGATGATGAT 2880
 QY 2879 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2938
 Db 2881 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
 QY 2939 AGTATTCATTTAAGAGAAATCCACACACACTGAGGAGCCGAGCTTGAT 2985
 Db 2941 AGTATTCATTTAAGAGAAATCCACACACACTGAGGAGCCGAGCTTGAT 2987

RESULT 3
 US-10-355-218-3
 ; Sequence 3, Application us/10355218
 ; Publication No. US20040009558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MORAS, DINO
 ; APPLICANT: RENAUD, JEAN-PAUL

APPLICANT: GRESCHIK, HOLGER
APPLICANT: WURZ, JEAN-MARIE
TITLE OF INVENTION: THE LIGAND BINDING POCKET PEPTIDE FRAGMENT OF THE
TITLE OF INVENTION: ESTROGEN-RELATED RECEPTOR 3 (ERR3) AND USES THEREOF
FILE REFERENCE: 0510-1060
CURRENT APPLICATION NUMBER: US/10/355,218
PRIOR FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: 60/352,551
PRIOR FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 3
LENGTH: 1377
TYPE: DNA
ORGANISM: Homo sapiens
US-10-355-218-3

Query Match 46.1%; Score 1377; DB 16; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 ATGAGTTCGTGAGACCTTGGCTTCTGATCTTTTCCCTGCACTAGAGAGAGCTT 185
DB 1 ATGAGTTCGTGAGACCTTGGCTTCTGATCTTTTCCCTGCACTAGAGAGAGCTT 60
QY 186 CTCTGCAAGATGTCAAAAGATGACACATTTGCACTGCTGCTTCATCAAG 245
DB 61 CTCTGCAAGATGTCAAAAGATGACACATTTGCACTGCTGCTTCATCAAG 120
QY 246 AGGAGACCTTCAGAGCCGCTCTCTGACGACAGACCTTCACACAGCCCTGCTGAC 305
DB 121 ACGGACCTTCAGAGCCGCTCTCTGACGACAGACCTTCACACAGCCCTGCTGAC 180
QY 306 TCTTCAGAGCCGCTTCAGAGCTACAGTCAACATGAGGCTTCAGAGCCGCTTCAG 365
DB 181 TCTTCAGAGCCGCTTCAGAGCTACAGTCAACATGAGGCTTCAGAGCCGCTTCAG 240
QY 366 TCGCACCCTCTCTACCTTCTGCTCTATCTCTGAGAGGTAGTGGCTCTGCAAGAACTG 425
DB 241 TCGCACCCTCTCTACCTTCTGCTCTATCTCTGAGAGGTAGTGGCTCTGCAAGAACTG 300
QY 426 TATGATGACTGCTCCAGACCACTTTGTAAGATCCCGACAGCAAGTGTGAATACATGCTC 485
DB 301 TATGATGACTGCTCCAGACCACTTTGTAAGATCCCGACAGCAAGTGTGAATACATGCTC 360
QY 486 AACTGATGAGCCCAAGAGACTGTGTTAGTGTGAGTGAATGCTCTGCTGCTGACATAT 545
DB 361 AACTGATGAGCCCAAGAGACTGTGTTAGTGTGAGTGAATGCTCTGCTGCTGACATAT 420
QY 546 GGGGTAGCATCATGTGAGACCTGCAAGGCACTTCTTCAGAGGACCAATTCAGGCAATATA 605
DB 421 GGGGTAGCATCATGTGAGACCTGCAAGGCACTTCTTCAGAGGACCAATTCAGGCAATATA 480
QY 606 GAATACAGCTGCTCCGACCAAGATGATGTAATACAAAGCCGACAGCTTAATCTGCG 665
DB 481 GAATACAGCTGCTCCGACCAAGATGATGTAATACAAAGCCGACAGCTTAATCTGCG 540
QY 666 CAGGCTGCGCGCTTCATGAAGTGTAAAGTGGAGCATGCTGTAAGAGAGGCTGCTCTT 725
DB 541 CAGGCTGCGCGCTTCATGAAGTGTAAAGTGGAGCATGCTGTAAGAGAGGCTGCTCTT 600
QY 726 GACAGAGTACGTGAGGTGCGGCAAGATCAAGCCGAGGATAGATGCGAGAAACAGCCCA 785
DB 601 GACAGAGTACGTGAGGTGCGGCAAGATCAAGCCGAGGATAGATGCGAGAAACAGCCCA 660
QY 786 TACCTGAACCTTCAGCTGCTGACGACGCAAAAGCCATATATACAAAGTGTCTACAT 845
DB 661 TACCTGAACCTTCAGCTGCTGACGACGCAAAAGCCATATATACAAAGTGTCTACAT 720
QY 846 TTGTTGTTGCTGAACCGGAGAGATCTATGCGATGCTGACCTTCTGCTGCTGCTGACAGT 905
DB 721 TTGTTGTTGCTGAACCGGAGAGATCTATGCGATGCTGACCTTCTGCTGCTGCTGACAGT 780

QY 906 GACATCAAGCCCTCACTACATGTTGATCTTGAGCCGACGAGAGTGTGTTATCATTT 965
DB 781 GACATCAAGCCCTCACTACATGTTGATCTTGAGCCGACGAGAGTGTGTTATCATTT 840
QY 966 GATGAGGCAAGATATTCAGAGCTTCTGACGCTGTCTCTGCGGACCAAGATGAGCTT 1025
DB 841 GATGAGGCAAGATATTCAGAGCTTCTGACGCTGTCTCTGCGGACCAAGATGAGCTT 900
QY 1026 CTGCAAGTGTGATGAGAAATTTGATCTTGTGTCTGTTATCCGCTCTCTTCAATTT 1085
DB 901 CTGCAAGTGTGATGAGAAATTTGATCTTGTGTCTGTTATCCGCTCTCTTCAATTT 960
QY 1086 GAGATGAATCTGTCTATGACAGATATATATATATGACGAGACCAAGTCCAAATATGCA 1145
DB 961 GAGATGAATCTGTCTATGACAGATATATATATATGACGAGACCAAGTCCAAATATGCA 1020
QY 1146 GGCCTTCTGATCTAAATATATGCTATCTGACGCTGTGTAAGAAATACAGAGCATGAG 1205
DB 1021 GGCCTTCTGATCTAAATATATGCTATCTGACGCTGTGTAAGAAATACAGAGCATGAG 1080
QY 1206 CTGAAAAAGAAATTTGTCACCTCTGAAAGCTATAGCTTTGCTTAATTCAGACTCCATG 1265
DB 1081 CTGAAAAAGAAATTTGTCACCTCTGAAAGCTATAGCTTTGCTTAATTCAGACTCCATG 1140
QY 1266 CACATGAAGATGTGAAAGCCGTTCAAGAGCTTCAGAGTGTCTTACATGAAGCGCTGAG 1325
DB 1141 CACATGAAGATGTGAAAGCCGTTCAAGAGCTTCAGAGTGTCTTACATGAAGCGCTGAG 1200
QY 1326 GATTATGAAGCTGACGACATGGAAGACCTCTGAGCTGCGCAAGTGTGATGCA 1385
DB 1201 GATTATGAAGCTGACGACATGGAAGACCTCTGAGCTGCGCAAGTGTGATGCA 1260
QY 1386 CTGCCACTCTGAGGACAGCTTACCAAGCCGTTGACGATTTTAAACATCAACTCA 1445
DB 1261 CTGCCACTCTGAGGACAGCTTACCAAGCCGTTGACGATTTTAAACATCAACTCA 1320
QY 1446 GAGGCAAGTCCCAAGTCCCAAACTTTTGGAAAAAGTGGAGCCCAAGTGTGCA 1502
DB 1321 GAGGCAAGTCCCAAGTCCCAAACTTTTGGAAAAAGTGGAGCCCAAGTGTGCA 1377

RESULT 4

US-10-054-841-1
; Sequence 1, Application US/10054841
; Publication No. US20030119100A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: 199993YCA
; CURRENT APPLICATION NUMBER: US/10/054,841
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: 09/487,379
; PRIOR FILING DATE: 2000-01-18
; PRIOR APPLICATION NUMBER: 09/141,000
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/078,633
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/062,902
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/057,090
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2807
; TYPE: DNA
; ORGANISM: Human
US-10-054-841-1

Query Match 24.0%; Score 715; DB 15; Length 2807;
Best Local Similarity 71.5%; Pred. No. 8,8e-178;
Matches 956; Conservative 0; Mismatches 375; Indels 6; Gaps 1;

```

QY 179 AGAGCTTCTGCAAGATGTCAAAACAAGATGACATTTGATTCAGCTGTTGCTCT 238
Db 934 AGGCTGTGTAACAGATGTCCTGGAGCAGAGGACTGGCTCCACTGGCTCTCT 993
QY 239 CATCAAGACGAACTTCCAGCTCCCTCTGACGAGCAGCTCAACCAACAGCC 298
Db 994 CATCAAGCTAGCGCTCCAGCTCCCTCTGAGGCAATAGTCCCTCAACCAAGCC 1053
QY 299 TGTGAGCTCTTCAAGCGCAGTGGAGCTACAGTTCAACCAATGAGCCATCAAGCG 358
Db 1054 CAGTGGCTGCTCCAGCGCAGCGGCTTGGCTGAGCTCCAGCAACCAAGCG 1113
QY 359 ACTTGACTCGCACTCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 418
Db 1114 TCTGAGCTCGCACCC-----CATGTTGACGCGCGCGGCTGGAGGCAACCATGCG 1167
QY 419 GAACTGTATGATGATGCTCTGACGACCATGTTGAAAGATCCCAAGCAAGTGTAA 478
Db 1168 CAAAGCTACGAGAGCTGTGCGACGCGCATATGAGAGCACTCGGCAATCAAGTGCAGTA 1227
QY 479 CATGCTCACTCGATGCGCCAAAGAGACTGTGTTAGTGTGTGATCATGCTCTGAGGTA 538
Db 1228 CATGCTCAAGGCCATCCCAAGCGCTGTGCTGTGTGCGGAGCAATGCTCTGAGTA 1287
QY 539 CCACTATGAGGCTAGCATCATATGTAAGCTGCAAGGCACTTCTTCAAGAGCAATTCAGG 598
Db 1288 CCACTAGCGGCTGCTCTCTGCGAGGCTTCAAGGCTCTTCAAGAGCAATTCAGAG 1347
QY 599 CAATATGATATACAGCTGCGCTGCGCAAGATGTAATGTAATTCAAAGCGCAGACTAA 658
Db 1348 GAACATTAAGTACAGCTGCGCGCGCACCAAGATGAGATGCAACCAAGCGAGCCCAA 1407
QY 659 ATCTGCGCAGCTTCCGCTTCAATGAAGTGTAAAGTGGCATGCTGTAAGAGGGGT 718
Db 1408 GTCTGCGCAGCGCTGCGCTTCAATGAATGCTCAAGTGGGATGCTGAAGAGGTGT 1467
QY 719 GCGCTTGAAGAGTACGTGAGAGTGGCGAAGATCAAGCAGAGATAGATGCGAGAA 778
Db 1468 GCGCTTGAAGAGTACGTGAGAGTGGCGCAAGATCAAGCAGAGTGTGAGAG 1527
QY 779 CAGCCCATACCTGTAACCTCAGCTGCTGTCAGCCAGCCAAAGCAATTAACAAGATTG 838
Db 1528 CAGCCCATACCTGTAACCTCAGCTGCTGTCAGCCATTAAGCAATTAACAAGATTG 1587
QY 839 CTCACATTTGTTGTTGCTGTAACCGAGATCTATGTCATGCTGCTGCTGCTGCTGCT 898
Db 1588 CTCACATTTGTTGTTGCTGTAACCGAGCTCTATGTCATGCTGCTGCTGCTGCTGCT 1647
QY 899 CGACATGATATCAAGCGCTCTACATGCTGTGACTGTGCTGCTGCTGCTGCTGCTGCT 958
Db 1648 TGAAGGAGATCAAGCGCTCTACATGCTGTGACTGTGCTGCTGCTGCTGCTGCTGCT 1707
QY 959 TATCATTTGATGAGGCGAAGCATATTCAGGCTCTTCCAGCGCTTCCCTGCGGAGCAAG 1018
Db 1708 CATCATTTGATGAGGCGAAGCATATTCAGGCTCTTCCAGCGCTTCCCTGCGGAGCAAG 1767
QY 1019 GAGCCTTCTGCAAGTCTTGTGATGTAATTTGATCTTGTGCTGCTGCTGCTGCTGCTGCT 1078
Db 1768 GAGCCTTCTGCAAGTCTTGTGATGTAATTTGATCTTGTGCTGCTGCTGCTGCTGCTGCT 1827
QY 1079 TTTATTTGAGATGAACTTGTCTATGCAAGCATATTAATTAATGAGCAAGCAAGTCCAA 1138
Db 1828 GGCCTTACGAGCAAGCTGTTGATGCTGAGCACTATCATGATGAGAGCAAGTCCCG 1887
QY 1139 ATTAGAGGCTCTTCTGATCTAATAATGCTATCCGAGCTGCTGTAAGAAATTAAGAG 1198
Db 1888 CCTGCGGGGCTGCTGAGCTCTACCGGCACTCTGAGCTGTAAGAGGTAAGAA 1947
QY 1199 CATGAAGCTGAAAAAAGAAATTTGTCACTCTCAAGCTATAGCTTGTCTAATTCAGA 1258
Db 1948 GCTCAAGGTGAGAAAGAGAGTGTGTGAGCTCAAGGCTTGGCTTGGCACTCCGA 2007

```

```

QY 1259 CTCATGCACTAGAAAGATGTTGAAGCCCTTCAGAACTTCAGATGTCCTTATCATGAAGC 1318
Db 2008 TTCACTGTAATCGAGATCTTAGAGGCTGTCTCAGAAAGCTTGACAGGCTGCTGCAAGAGC 2067
QY 1319 GCTGAGATTTATGAAGCTGAGCAGACATGAAAGACCTGCTGAGCTGCGAAAGTCT 1378
Db 2068 ACTGAGAGTATACAGAGCTGAGCCAGCGCCATGAGAGGCTCTGAGAGAGGCGCAAGTCT 2127
QY 1379 GATGACATGCTCACTCTGAGGAGAGCTCTTACCAAGGCGCTGACATTTCTACAACT 1438
Db 2128 GCTACACTGCGCTGCTGCGGAGAGCGCGCAAGGCGCTGACACTTCTATAGCGT 2187
QY 1439 CAACTGAAAGCAAGTCCCAATGCACAACTTTTGTGAATGTTGAGGCTCAAGT 1498
Db 2188 CAACTGAGGCGCAAGTCCCAATGCACAACTTCTGAGATGCTGAGGCGCAAGC 2247
QY 1499 CTGACTAAAGCTCCCT 1515
Db 2248 CTGAGCGAGGCTGACT 2264

RESULT 5
US-10-101-510-87
; Sequence 87, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: MAN JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 87
; LENGTH: 2153
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-87

Query Match 23.3%; Score 695; DB 15; Length 2153;
Best Local Similarity 69.7%; Pred. No. 1,4e-172;
Matches 972; Conservative 0; Mismatches 415; Indels 8; Gaps 2;

QY 108 AATGTTGCGGTGCAATGATTCGGTAGAATTTGCTTCTGATCTTTTCTCTG 167
Db 15 AATGCTAATAAGGAGCTGATGAGAGTGTCCGAATCTGCAATCCCGAGCCCCCTGAGTAC 74
QY 168 CACTAGAGGAGAGAGCTTCTCTGCAAGATGTCAAACAAAGATGAGCATTTGATTCAGC 227
Db 75 CACACACAGTA--GTTTGTGAACGAAATGTCGAGAAACAGGACCTGGGCTTAGC 132
QY 228 TGTTCCTCTTCAATCAAGACGAACTTCCAGGCGCAAGCTCCCTGACGAGAGGTCAAC 287
Db 133 TGGGCTCTCTTATCAAGAGGAGCATCTAGCCCATCTCGGCAATGATGCTTAC 192
QY 288 CACCAAGCCCTGTGCTCTTCAAGCGCAAGTGGAGCTTACAGTTCAACATGAATGGC 347
Db 193 CACCAAGCCCGAGCGGCTGTGAGCGCAAGGCTGGCTTTGGCAATGGGCTGGGAGC 252
QY 348 CATCAAGAGAGCTGAGCTGCGCACTCTCTACCTTCTGCTCTATCTCTGAGAGTACT 407
Db 253 CAGCCCAAGGCTGAGCTCTCGCTATGTTGCGAGGTC-----GGGAGCTGGAGGC 306
QY 408 GGGCTGTGAGAACTGTATGATGATGCTGCTCAAGCACTTGTGAAGTCCCGAGAC 467
Db 307 AACCGGTGCGCAAGAGCTACAGAGACTGTACTAGGAGATATGAGAGACTGGCCATC 366
QY 468 AAGTGTGAATCATGCTCACTGATGCCCAAGAGACTGTGTTAGTGTGTGATCATC 527
Db 367 AAGTGTGAATCATGCTTAAAGCCATCCCAAGCGCTGTGCTGTGTGCGGAGCATT 426

```


QY 888 CTTACTGTCCCGACAGTACATCAAAAGCCCTCACTACATGTGNGACTTGGCCGACCGA 947
DB 787 GACGATGTGCTGGAAGGGATATCAAGGCCCTGACACTCTCTGTGACTTGGCAGATGG 846
QY 948 GAGTGTGTGTTATCATTTGATGAGGCGAAGCATATTCAGGCTTCTCCAGCTGTCCCTG 1007
DB 847 GAGCTGTGTGTTCCATTTAGCTGGGCGACAGCATCCAGGTTTCTCCAACTGACACTC 906
QY 1008 GCGGACCAAGATGAGCCCTTTCAGAGTGTGATGAGAAATTTGATCTTGGTGTGTA 1067
DB 907 GGGGACCAAGATGAGCCCTGTCAGAGTGTGATGAGATCCATCTCGGAGATGAG 966
QY 1068 TACCGGTCTCTTTCATTTAGAGTAACTTGTATGAGACATTAATATGAGACAA 1127
DB 967 TACCGGTCTCTTTCATTTAGAGTAACTTGTATGAGACATTAATATGAGACAA 1026
QY 1128 GACCAATCCAAATATGAGAGCCCTTGTATGATTAATATGATCTGACAGCTGTAAAG 1187
DB 1027 GACCAATCCAGCTGAGGCTGTGAGGCTGTGAGCTTACCGAGCATCTTGCAGCTGTAGGC 1086
QY 1188 AATATCAAGAGCATGAGAGCTGAGAAAGAAATTTGTCAACCTCAAGCTATAGCTCTT 1247
DB 1087 AGGTACAGAGAGCTCAAGGTGAGAGAGAGATTTGTATGATCTCAAGCCCTGAGCCCTT 1146
QY 1248 GCTAATTCAGACTTCATGACATAGAGATGTTGAAAGCCGTTGAGAGCTTCAGAGATGTC 1307
DB 1147 GCGAATTCAGATTAATGATGATGAGAGACCTGAGAGCTGTGAGAGCTTCAGAGATGTC 1206
QY 1308 TTACATGAGAGCTGTGAGAGATTTATGAGAGCTGCGACACATGAGAAACCTCTGTGAGCT 1367
DB 1207 CTGATATGAGAGCTGTGAGAGATTTATGAGAGCTGAGAGCTGAGAGAGAGAGAGAG 1266
QY 1368 GCGAAGTGTGATGACATGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAT 1427
DB 1267 GCGAAGTGTGATGACATGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAT 1326
QY 1428 TTCTAACAATCAAACTAG 1487
DB 1327 TTCTAACAATCAAACTAG 1386
QY 1488 GAGGCCAAGGTCTGA 1502
DB 1387 GAGGCCAAGGTCTGA 1401

RESULT 7

US-10-054-841-8/c
Sequence 8, Application US/10054841
Publication No. US20030119100A1
GENERAL INFORMATION:
APPLICANT: Chen, Fang
TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
FILE REFERENCE: 19999YCA
CURRENT APPLICATION NUMBER: US/10/054,841
PRIOR FILING DATE: 2002-01-23
PRIOR APPLICATION NUMBER: 09/487,379
PRIOR FILING DATE: 2000-01-18
PRIOR APPLICATION NUMBER: 09/141,000
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/078,633
PRIOR FILING DATE: 1998-03-19
PRIOR APPLICATION NUMBER: 60/062,902
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/057,090
PRIOR FILING DATE: 1997-08-27
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 622
TYPE: DNA
ORGANISM: Human

FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(622)
OTHER INFORMATION: n = A,T,C or G
US-10-054-841-8

Query Match 14.3%; Score 428; DB 15; Length 622;
Best Local Similarity 93.9%; Pred. No. 2.9e-102;
Matches 495; Conservative 0; Mismatches 27; Indels 5; Gaps 5;

QY 25 CGGCTTGCTCACTAGAGAAATTTGTGTTAATTCACATGCTGCTGTCAGAGAACTTTG 84
DB 546 CAGTTGATCCCNANMAGAGCATTTTGTATGTCCTGTCTGTGCAAGG-AACTTTG 489
QY 85 ATTATATGCTGGGATGACAAATTAATGTTGCGCGGTGGACATGATTC-GGTGACTT 143
DB 488 ATTATATGCTGGGATGACAAATTAATGTTGCGCGGTGGACATGATTCGGGAAAGACTT 429
QY 144 TGCTTCTGTAATCTTTTCTCTGCACTAC-GAGGAAAGCTTCTGTCAGAGATGT-CAA 201
DB 428 TGCTTCTGTAATCTTTTCTCTGCACTACGAGGAGGAAAGCTTCTGTCAGAGATGTCAA 369
QY 202 ACAAGATGACACATTTGATTCAGCTGTTGCTCTTATCAAGACGAACTTCCAGCC 261
DB 368 ACAAGATGACACATTTGATTCAGCTGTTGCTCTTATCAAGACGAACTTCCAGCC 309
QY 262 CAGCTTCTGAGAGAGAGGTCAACCAAGCCCTGTGGTGTCTTTCAGAGCCAGTG 321
DB 308 CAGCTTCTGAGAGAGAGGTCAACCAAGCCCTGTGGTGTCTTTCAGAGCCAGTG 249
QY 322 GAGCTACAGTTCAACCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 381
DB 248 GAGCTACAGTTCAACCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 189
QY 382 CTCTGCTCTTATCTGAGAGTGTGAGGCTGTGAGAGAACTGATGATGATGATGATGATG 441
DB 188 CTCTGCTCTTATCTGAGAGTGTGAGGCTGTGAGAGAACTGATGATGATGATGATGATG 129
QY 442 GCACCATGTTGAAGATCCCAAGCAAGTGTGATGATGATGATGATGATGATGATGATG 501
DB 128 GCACCATGTTGAAGATCCCAAGCAAGTGTGATGATGATGATGATGATGATGATGATG 69
QY 502 GACTGATTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 548
DB 68 GACTGATTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 22

RESULT 8

US-09-962-436-268
Sequence 268, Application US/09962436
Patent No. US20020081301A1
GENERAL INFORMATION:
APPLICANT: Soppet, Daniel
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-75
CURRENT APPLICATION NUMBER: US/09/962,436
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/60/235,082
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/234,924
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 568
SOFTWARE: PatentIn version 3.0
SEQ ID NO 268
LENGTH: 2218
TYPE: DNA
ORGANISM: Homo sapiens
US-09-962-436-268

Query Match 14.2%; Score 424.4; DB 9; Length 2218;
Best Local Similarity 64.7%; Pred. No. 6.4e-101;
Matches 682; Conservative 0; Mismatches 336; Indels 36; Gaps 2;

QY	1117	TAATGGAAGAAACACAGTCCAAATTAGAGAGCCTTTGTAATCTAAATAAGCTACTCTGC	1176
Dp	1046	TCTCGATGAAGAAGGGGCGACGGGCACTGTGCTCGGGGAACTGGGGCTCCCTGCTGC	1105
QY	1177	AGCTGTAAAGAAATACAGAGCAATGAAGCTGGAAGAAAGAAATTTGTACCCCTCAAG	1236
Dp	1106	AACTAGTGTGGGGGTGTGAGGCCCTCGGGCTGAGACGAGAGAGTAATTTTACTTAAGG	1165
QY	1237	CTATAGCTCTTGTCTAATTCAGACTTCATGACATAGAAGATGTTGAACCGTTCAGAAGC	1296
Dp	1166	CCTTGGCCCTTGCCAAATTCAGACTCTGTGGCATTGCAATGCAAGATGCGAGCTGTGAGGAGC	1225
QY	1297	TTTCAGATGTTCATCATAGAACCGCTGCAAGATTATGAAGCTGCGCAG-----C	1344
Dp	1226	TGCGAAGAGCTCTGTGACGAGGCCCTGCTGAGATATGAAGCCGGGCGGGCTGGCCCGGAG	1285
QY	1345	ACATGGAAGACCTTCGTGAGCTGGCAAGATGCTGATGACCTGCACTCCTGAGGGCAGA	1404
Dp	1286	GGGTTGTGAGCGGGCGGGCGGGCGGAGCTGTGCTCAGCCTACCGTCTCTTCGCCAGA	1345
QY	1405	CCTTACCAAGGCCGTGCAGATTTTCAACAATCAACTGAAGGCAAGATCCCAATGC	1464
Dp	1346	CAGCGGCAAAAGTCTGGCCATTTCATATGGGGTGAAGCTGGAAGGCAAGTGGCCATGC	1405
QY	1465	ACAAACTTTTTTGGAAATGTTGAGAGCCAAAGT	1498
Dp	1406	ACAAGCTGTTCTTGAGATGCTCGAGGCCAATGAT	1439

RESULT 10

US-10-037-270-454
; Sequence 454, Application US/10037270
; Publication No. US20030104529A1

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyang
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aйдong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dairui
APPLICANT: Wang, Zhimei
APPLICANT: Tillinghast, John
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529a1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CTP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 454
LENGTH: 2539
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (97)..(1755)
US-10-037-270-454

Query Match	14.1%;	Score 421.4;	DB 15;	Length 2539;
Best Local Similarity	64.5%;	Pred. No. 4.4e-100;		

Matches 682; Conservative 0; Mismatches 336; Indels 39; Gaps 2;

QY	481	TGCTCAAC	CGTCAATGCC	CCAAAGAC	CTGTGTTT	AGTGTGTG	GTACAT	TCGCTT	CTGGATCC	540					
Db	692	TGCTGAGCT	CCCTGCCC	CAAGCCCT	CTGTGCTGTG	TGGGGAC	GTGGCT	CCGACT	CC	751					
QY	541	ACTATGGG	TGATCAT	TGTGAA	GCTGTG	CAAGCCAT	TTCTTCA	GAAGCA	CAATTC	GAAGCA	600				
Db	752	ACTATGGT	TGCTCAT	CTGTGT	GAAGCT	CTTCTTCA	GAAGCA	CAATTC	GAAGCA	811					
QY	601	ATATGAA	TACAGCT	TGCTCC	TGCACGA	ATGAAT	TGTGAAT	CA	CAAGCG	CAGAC	660				
Db	812	GCATGAG	TACAGCT	TCGGGCT	CCAA	CGATGT	GAATCA	CCAA	CGGAG	CCAA	871				
QY	661	CTGCGAGG	CTTGCGCTT	CATGA	AGTGT	TAAAGT	GGGCA	TG	TGAAAG	AGGGGT	720				
Db	872	CCTGCAAG	CGCTCGCGCTT	CACCA	AGTGT	CGGAT	GGGAT	CTTCA	GAAG	GGAG	931				
QY	721	GTCTTGA	CAAGTA	CGTGA	AGT	CGGACGA	AGTCA	CAAGCG	CAGAT	GAATGCG	GAACA	780			
Db	932	GCTTGAC	CGCGCT	CCGGGAT	TGGGCGG	CAGA	AGTCA	CAAGCG	CGGAG	CCCAAG	991				
QY	781	GCCCAT	TACCTGA	CCCTCAG	CTGT	CAAGC	-----	-----	-----	AG	813				
Db	992	TGCCCTT	CCCGGCCCCCT	CTCCCTG	TGGACCTT	GGACGT	CGCTG	GAAG	CCCCCG	GAAGA	1051				
QY	814	CCAAAAG	CCCATAT	TAA	CAAGAT	TGTCT	CACAT	TTGTT	GGTGG	CTGAAC	CCGAGAA	873			
Db	1052	CAGACGCC	CCAGTGA	ATGCAC	TGTCT	CATCTGT	GTGTG	ATGAG	CTTGA	AGGCT	CT	1111			
QY	874	ATGCGAT	GCTGAC	CTCTCT	GTCCCC	CACAGT	GACAT	TAAAG	CCCTCA	CTAC	ACTGT	933			
Db	1112	ATGCCAT	CCCTGA	CCCCCG	CAGCCCT	GTATGG	GCACCT	CCAGCG	TGCTAC	CCCTGT	GTG	1171			
QY	934	ACTTGG	CCGAC	CCGAG	AGTTGG	TGTAT	CAT	TGAT	TGGG	CGCAAG	ATTTT	COAGGCT	993		
Db	1172	ACCTCTT	GACCG	AGAT	TGTGT	GCAC	CACT	GA	CTTGG	CCCAAG	CAAT	CC	1231		
QY	994	CCAGCT	GTCCCT	TGCGG	ACAGAT	GAAG	CCCTT	TG	CAGAG	TGCTT	GTGAT	GTGA	1053		
Db	1232	CATGCT	GTGCTGT	CTGAC	CAGAT	TGT	CAGT	ACT	TG	CAGAG	CGTGT	GGAT	1291		
QY	1054	TCTTGT	GTGTGT	ATACCG	GTCTCTT	CATTT	TGA	AGAT	TGAAT	CTTGT	CTAT	CAAC	AT	1113	
Db	1292	TGCTGG	GTGTGG	CCCAAG	CGCTCA	CTGTGC	AC	CTG	CAC	GTG	CAGAT	TGA	CT	1351	
QY	1114	ATTATAT	GACGA	AGAC	CCAGT	CCAAAT	TAG	GAGG	CCCTT	CTT	GAAT	CTAAAT	AT	1173	
Db	1352	TAGTCT	TGGA	TGAAG	GGGGG	CA	CGGG	AGT	TGG	CT	TGGGG	AACT	GT	1411	
QY	1174	TGCAGCT	GTGTAA	GAATAT	CAAG	AGCAT	GAAG	CTG	GA	AAAA	AGAA	TTT	GT	1233	
Db	1412	TGCAACT	AGTGTGG	CGGCTGT	GACG	CCCTG	CGCTG	AG	CGGAG	CGAG	AGAT	TTT	CT	1471	
QY	1214	AAGCTAT	AGCTT	TGCTTA	TTCA	CAAT	CCAA	TG	CAAT	TA	GAAG	AT	GT	1239	
Db	1472	AGGCTT	GTGGCTT	GTCCAT	TTCA	CAAT	CTGT	TG	CAAT	TA	GAAG	AT	GT	1531	
QY	1294	AGCTT	CAGAGT	GTCTTA	CAT	GAAC	CGCTG	CAG	ATTA	TGA	AGCT	GTGG	CCAG	-----	1343
Db	1532	AGTGT	CGAGAG	GTCTGT	CGAC	AGG	CCCTGT	G	AGT	TGA	AG	CGG	CG	GTGG	1591
QY	1344	--	CACAT	TGA	AGAC	CCCTGT	TG	GAAC	CTGT	GAAC	CTGT	GAAC	CTGT	GAAC	1401
Db	1592	GAGGG	GTGTGT	GAAC	CGGCGG	CGGCGG	CAAG	CGCTGT	CA	CACT	CA	CGCT	CT	CGG	1651
QY	1492	AGACCT	CTAC	CAAG	CGCGT	GTG	CAGAT	TTT	CA	CAAT	CA	CAAT	CA	CAAT	1461
Db	1652	AGAC	AGGCGG	GAAG	GTGT	GTGG	CCAT	TTT	CT	T	GGGG	GT	GAAG	CT	1711
QY	1462	TGCA	CAAACT	TTTTT	TG	GAAT	TTT	TG	AG	CG	CAAG	AGT	1498		
Db	1712	TGCA	CAAGCT	TTT	CTT	TG	GAAT	TTT	TG	AG	CG	CAAG	AGT	1748	


```

RESULT 11
US-10-117-722-454
; Sequence 454, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Dhanraj, Radhaje T.
; TITLE OF INVENTION: No. US20030219744A1e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_files Version 1.0
; SEQ ID NO 454
; LENGTH: 2539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97) .. (1755)
US-10-117-722-454

Query Match      14.1%; Score 421.4; DB 16; Length 2539;
Best Local Similarity 64.5%; Pred. No. 4.4e-100;
Matches 682; Conservative 0; Mismatches 336; Indels 39; Gaps 2;

481 TGCCTCACTCGATGAGCCCAAGAGACGTGTGTTAGTGTGTGATGATCGCTTGTGGGTACC 540
    |||||
692 TGCCTCACTCGCTCCGCCCAAGCGCTTGCCTGCTGTGCGGAGCGTGGCTCCGGCTTACC 751
    |||||
541 ACTATGGGGTAGCATATGATGAAAGCTGCAAGGCAATTTTCAAGAGCAATTCAGAGCA 600
    |||||
752 ACTATGGTGGGATCTGTGAGGCTGCAAGGCTTCTTCAAGAGCAATTCAGAGCA 811
    |||||
601 AATATAAATACAGCTGCTCCCTGCGACGAATGATGAAATCAAGAGCGAGCACTAAT 660
    |||||
812 GCATCGAGTACAGCTCTCGGCTTCAAGAGTGTGATCAAGAGCGAGCAAGG 871
    |||||
661 CCGCAGAGCTTGCCTTCATGAGTGTAAAGTGGGCATGCTGAAAGAGAGGGGTGC 720
    |||||
872 CCGCAGAGCTTGCCTTCATGAGTGTAAAGTGGGCATGCTGAAAGAGAGGGGTGC 931
    |||||
721 GTCTTACAGAGTACGTGAGGTGCGGAGAGTCAAGCGCAGAGATGATGCGGAGAAC 780
    |||||
932 GCTTGACGCGCGTCCGCGGTGCGGAGAGTCAAGCGCGCGAGGTGAGCCAC 991
    |||||
781 GCCCATACCTGAACCTCGAGCTTCAAGC-----AG 813
    |||||
992 TGCCCTTCCCGGGGCTTCCCTGCTGGGGCTTGGCAGTGCCTGAGAGGCCCCCGAGAG 1051
    |||||
814 CCAAAAAGCATATTAACAAGATGTCTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 873
    |||||
1052 CACAGAGCCCAAGAGATGACCTGTGTCTCATCTGCTGGTGTGAGCTGAGAGGCTCT 1111
    |||||
874 ATGCCATGCTTGAACCTTACTGTCCCGACAGTACATCAAGAGCCCTTACTACACTGTGTG 933
    |||||
1112 ATCCATGCTTGAACCCCGACAGGCTGATGAGGCACTCCCAAGCGGTGCTTACCTCTGTG 1171
    |||||
934 ACTTGGCCGACGAGAGATGTTGTTATCATTTGGATGGGCGAAGCATATTCAGAGCTTCT 993
    |||||
1172 ACCTCTTGAACGAGATGTGTGTCACTACAGCTGGGCGAGAGCATTCAGAGCTTCT 1231
    |||||

```

```

994 CCAGCTGTCCCTTGGCGGACCCAGATGAGCTTTCTGAGAGTGTGATGGAATTTGA 1053
    |||||
1232 CATGCTGTGCTGTCTGTACAGATGTCACTGACAGAGGTGTGATGAGAGGTGCTGG 1291
    |||||
1054 TCTTGGTGTGTATACCGGTCTCTTTCATTTGAGAGTGAATTTGTCTATGACAGATT 1113
    |||||
1292 TGTGGGTGTGGCCAGGCTCTCATCTGCTACATGACAGATGAGTGGCTTCTGCTAGGACT 1351
    |||||
1114 ATATATATGAGAGAGACCAAGTCCAAATTAGAGGCTTCTGTGATTAATATATATATATAT 1173
    |||||
1352 TAGTCTGTGATGAGAGAGGAGGACAGGAGCTGGCTGGGGAATCTGGAGGCTGCTTGC 1411
    |||||
1174 TGCAGCTGTATAAGAAATACAAAGCATGAACTGAGAAAGAAATTTGTACACCTTCA 1233
    |||||
1412 TGCACATGTGTGGGGGTGACAGGCTTGGCGCTGGAACGAGAGATGATTTACTTAA 1471
    |||||
1234 AAGCTATAGCTCTTCTATTTACAGTCACTGATGATGATGATGATGATGATGATGATGATGAT 1293
    |||||
1472 AGGCTTGGCTTGTCCCAATTCAGCTCTGTGATGATGATGATGATGATGATGATGATGATGAT 1531
    |||||
1234 AGCTTCAGAGATGTCTTACATGAGAGGCTGACAGATTTAGAGCTGAGC----- 1343
    |||||
1532 AGCTCGAGAGCTGTGACAGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1591
    |||||
1344 --CATATGAGAGCTCTGTGAGCTGTGAGATGATGATGATGATGATGATGATGATGATGATGAT 1401
    |||||
1592 GAGGGGTGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1651
    |||||
1402 AGACCTTACCAAGGCGGTCAGCATTTCTACATCAATCAATCAATCAATCAATCAATCAATCAAT 1461
    |||||
1652 AGACAGCGGCGCAAGTGTGCGGCTTCTATGAGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1711
    |||||
1462 TGCACAACTTTTGTGAAATGTTGAGGCGCAAGT 1498
    |||||
1712 TGCACAGCTGTCTTGTGAGATGCTGAGGCGCATGAT 1748
    |||||

RESULT 12
US-10-297-653-2
; Sequence 2, Application US/10297653
; Publication No. US20030215829A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: CHINN, Anna M.
; APPLICANT: LU, Yan
; APPLICANT: TRIBOULEY, Catherine M.
; TITLE OF INVENTION: NUCLEAR HORMONE RECEPTORS
; FILE REFERENCE: PI-0119 USN
; CURRENT APPLICATION NUMBER: US/10/297,653
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/209,863
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 4250
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030215829A1 7475568CB1
US-10-297-653-2

Query Match      14.1%; Score 421.4; DB 16; Length 4250;
Best Local Similarity 64.5%; Pred. No. 6.4e-100;
Matches 682; Conservative 0; Mismatches 336; Indels 39; Gaps 2;

481 TGCCTCACTCGATGAGCCCAAGAGACTGTGTTAGTGTGTGATGATCGCTTGTGGGTACC 540
    |||||
398 TGCCTCACTCGCTCCGCCCAAGCGCTTGCCTGCTGTGTGCGGAGCGTCCGGCTTACC 457
    |||||
541 ACTATGGGGTAGCATATGATGAGGCTGCAAGGCAATTTCAAGAGCAATTCAGAGCA 600
    |||||

```

```

Db 458 ACTATGCTGTGGCACTCTGTGAGGCGCTGCAGAGCTTCTTCAAGAGGACATCCAGGGGA 517
QY 601 ATATGAATTAAGCTGCCCTGCGACAGATGATGGAATCAAAAGGCGAGACTAAT 660
Db 518 GCATGAGTACAGCTGTCTCGGCTTCAACGAGTGTGATCAACCAAGGCGAGACCCAGG 577
QY 661 CCTGCGAGGCTTGCCTGCTTCAATGAAGTGTAAAGTGGGCACTGTGAAGAGGGGGTGC 720
Db 578 COTGCGAAGGCTCGCGCTTCAACAGTGCCTGCGGTGGGATGCTTCAAGAGGGAGTGC 637
QY 721 GTCTTGAAGAGTACGTGAGAGTGGCGCAAGATACAGCGCAAGATGATGATGCGAGAA 780
Db 638 GCTTGAGACCGGCTCGGAGTGGCGGAGAGTACAAAGCGCGGCGAGGTGAGCCAC 697
QY 781 GCCCATACCTGAACCTCAGCTGGTTAGCC-----AG 813
Db 698 TGCCCTTCCGGGGCCCTTCCCTGCTGGGGCCCTGGCAGTGGTGGAGGCCCCCGAAGA 757
QY 814 CCAAAAAGCCATATTAACAAGATTGTCTCAATTGTTGGTGGCTGAACCGGAGAAATCT 873
Db 758 CAGAGCCCCAGTGAATCACTGGTGTCTCATCTGTGTGTGAGCTGAGAACTCT 817
QY 874 ATGCCATGCTGACCTCTACTGTCCCCCAGAGTGAATCAAGCCCTCACTNACTGTGTG 933
Db 818 ATGCCATGCTGACCCCGCAGGCGCTGATGGGCACTCCAGCGTGTACTACCTCTGTGT 877
QY 934 ACTTGGCGCAGCAGAGTGTGGTGTATCATTTGATGGGCGCAAGCATTTCCAGGCTTCT 993
Db 878 ACCTCTTGAACCGAGAGATTGTGTCACTACATAGCTGGGCCAAGAGCATCCAGGCTTCT 937
QY 994 CCAGCGTGTCCCTGCGCGAGCAGATGAGCCCTTCTGCAAGTGTCTGATGAGAAATTTGA 1053
Db 938 CATGCGTGTGCTGTCTGACAGATGTCACTGACAGAGCGTGTGATGAGAGTGTCTGG 997
QY 1054 TCCCTTGTGTGTATACCGGTCTCTTTCATTTGAGATGAATCTGTCTATGCAAGCAT 1113
Db 998 TGTGGGTGTGGCCAGCGCTCACTGCGACATGCAAGATGAGCTGCGCTTCCCTGAGACT 1057
QY 1114 ATATATGAGAGAGACAGTCCAAATTAGAGGCGCTTCTGATTAATTAATGCTAATCC 1173
Db 1058 TAGTCTTGGATGAAGAGAGGGGCGACGGGAGCTGGGCGGAACTGGGGGCTGCCCTGC 1117
QY 1174 TGCAGCTGTGAAGAAATACAAAGACATGAGCTGTGAAAAAGAAATTTGTCACTCTCA 1233
Db 1118 TGCAACTAGTGGCGGCGCTGAGGCGCTGCGGTGAGGCGAGAGAGATGTCTACTAA 1177
QY 1234 AAGCTATAGCTTGTCTAATTCAACTTCATGCACTGACATGAGAAATGTTGAAGCCGTTCA 1293
Db 1178 AAGCTTGGCCCTTGCATATTCAGATCTGTGCACTGCAAGATGCGAGGCTGTGAGC 1237
QY 1294 AGCTTCAAGATGTCTTCAATGAAGCGCTGCAAGATTAAGATGAGCTGGCGAG----- 1343
Db 1238 AGCTGCGAAGAGCTCTGCAAGAGCCCTGCTGAGATTAAGGCGGCGGCGCTGGCCCCG 1297
QY 1344 --CACTGAGAAAGCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1401
Db 1298 GAGGGGCTGTGAGGCGGCGGCGGCGGCGGAGGCTGTGCTCAAGCTTCCCTGCGCC 1357
QY 1402 AGAGCTTACCAAGCGCGTGCAGATTTCTTCAACATCAACTTGAAGGCAAGTCCCA 1461
Db 1358 AGAGAGGCGGCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1417
QY 1462 TGCACAAACTTTTGTGAAATGTTGAGGCGCAAGT 1498
Db 1418 TGCACAAAGCTGTCTTGTGAGATGCTGAGGCGCATGAT 1454

```

RESULT 13
 US-10-342-887-1041
 ; Sequence 1041, Application US/10342887
 ; Publication No. US20040058340A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van der Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT FILING DATE: 2003-01-15
; PRIOR FILING DATE: 2001-06-18
; PRIOR FILING DATE: 2001-06-18
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1041
; LENGTH: 2402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1041

```

```

Query Match 13.5% Score 401.6; DB 13; Length 2402;
Best Local Similarity 64.5%; Pred. No. 7.4e-95;
Matches 684; Conservative 0; Mismatches 334; Indels 42; Gaps 4;

```

```

QY 481 TGCTCACTGATCCCAAGAGACTGTGTATGATGTGTGATGATGATGATGATGATGATGAT 540
Db 563 TGCTCACTGATCCCAAGAGACTGTGTATGATGTGTGATGATGATGATGATGATGATGAT 622
QY 541 ACTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 623 ACTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 682
QY 601 ATATGAATTAAGCTGCCCTGCGACAGATGATGATGATGATGATGATGATGATGATGAT 660
Db 683 GCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 742
QY 661 CCTGCGAGGCTTGGCGCTTCAATGAAGTGTAAAGTGGGCGATGCTGAAGAGGGTGC 720
Db 743 COTGCGAAGGCTTGGCGCTTCAACAGTGTGCTGCGGAGTGTGAGAGTGTGCTCAAGAGGAGTGC 802
QY 721 GTCTTGAAGAGTACGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 780
Db 803 GCTTGAGACCGGCTCGGAGTGGGCGGCGGAGAGTACAAAGCGCGGCGAGGTGAGCCAC 862
QY 781 GCCCATACCTGAACCTCAGCTGGTTAGCC-----AG 813
Db 863 TGCCCTTCCGGGGCCCTTCCCTGCTGGGGCCCTGGCAGTGGTGGAGGCCCCCGAAGA 922
QY 814 CCAAAAAGCCATATTAACAAGATTGTCTCACTTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 873
Db 923 CAGAGCGCCAGAGAAATGACTGCTGTCTCATGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 982
QY 874 ATGCCATGCTGACCTTACTGTCCCCGAGTGAATCAAAAGCCCTCACTNACTGTGTG 933
Db 983 ATGCCATGCTGACCCCGAGGCGCTGATGGGCACTCCAGCGGTGACTAACCCTGTGTG 1042
QY 934 ACTTGGCGCAGCAGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993
Db 1043 ACCTCTTGAACCGAGATGTGTCTCACTCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1102
QY 994 CCAAGCTGTCCCTGCGCGAGCAGATGAGCCCTTCTGCAAGTGTCTTGTGATGAGAAATTTGA 1053
Db 1103 CATGCGTGTGCTGTCTGACCAAGTGTCACTGACAGAGCGTGTGATGAGAGTGTCTGG 1162
QY 1054 TCCCTTGTGTGTATACCGGTCTCTTTCATTTGAGATGAATCTGTCTATGCAAGCAT 1113
Db 1163 TGTGTGGTGTGGCCAGCGCTCACTGCACTGCAAGATGAGTGTGCTGTGAGACT 1222

```


; CURRENT APPLICATION NUMBER: US/10/211,239
 ; CURRENT FILING DATE: 2002-08-05
 ; PRIOR APPLICATION NUMBER: US 60/315,775
 ; PRIOR FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: EP 01119003.0
 ; PRIOR FILING DATE: 2001-08-06
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 2402
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-211-239-7

Query Match 13.5%; Score 401.6; DB 15; Length 2402;
 Best Local Similarity 64.5%; Pred. No. 7,4e-95;
 Matches 684; Conservative 0; Mismatches 334; Indels 42; Gaps 4;

QY 481 TGTCTCACTTCATGCGCCAAAGACTGTGTGTTAGTGTGTGACATCGCTTCTGGGTAAC 540
 Db 563 TGTCTAGCTCCCTGCCCAAGGCGCTCTGCTGTGTGTGGGAGCGTGGCTCCGGCTTACC 622
 QY 541 ACTATGGGGTAGCATCATGTGAAGCGCTGCAAGCGATTTCTTAAGAGCAATTCAAGGCA 600
 Db 623 ACTATGGTGTGCAATCTGTGAGGCTTCAAAAGCCTTCTTCAAGAGCAATCCAGGGGA 682
 QY 601 ATATAGATACAGCTGCCCTGCCAAGATGATGTAATCAAAAGCGCAGAGCTAAT 660
 Db 683 GCATCGAGTACAGCTGTCCGGCTCCAAAGGTGAGATACCAAGCGGAGAGCCAGAGG 742
 QY 661 CTTGCCAGGCTTGGCGCTTCAATGAGTGTAAAAAGTGGGCAATGCTGAAGAGGGGTGC 720
 Db 743 CTTGCCAGGCTTGGCGCTTCAACAAAGTGTCCGGGTGGGCAATGCTCAAGAGGAGTGC 802
 QY 721 GTCTTGACAGAGTGTGAGAGTGGGCGAGAGTCAAGCGCAGATAGATCGGAGAA 780
 Db 803 GCTTGACCGGCTGCGGGGTGGGCGGAGAGTCAAGCGGCGGAGGTGAGACCCAC 862
 QY 781 GCCCATACCTGAAACCTCAGCTGTGACCC-----AG 813
 Db 863 TGCCCTTCCGGGGGCCCTTCCCTGTGCGGCCCTGGCAGTGGAGAGCCCGGAGAA 922
 QY 814 CCAAAAAGCCATATACAAAGTTGTCTCAATTTGTGTGGTGAACCGGAGAAATCT 873
 Db 923 CAGCAGCCCGAGTAATGCACTGTGTGTCTCATCTGTGTGTGAGCTGAGAAAGCTCT 982
 QY 874 ATGCCATGCTGACCCCTAATGTCCCGCAGATGACATCAAGCCCTCACTACATGTGTG 933
 Db 983 ATGCCATGCTGACCCCGCAGGCGCTTATGAGCACTCCAGCGGTAGCCCTCTGTG 1042
 QY 934 ACTTGGCCGACCGAGAGTGTGTGTATTCATTTGATGGGCGAAGCATATTCAGAGCTTCT 993
 Db 1043 ACCTCTTTGACCGAGATTTGTGTCAACATCACTGGGCCAAGAGCATCCAGAGCTTCT 1102
 QY 994 CCAAGCTGTCCCTTGGCGGACCATGAGCTTCTGCAAGTGTGTGAGTGAATTTGA 1053
 Db 1103 CATGCTGTGTGTGTGACAGATGTCACTGACAGAGGTGTGATGAGGTGTGCTGG 1162
 QY 1054 TCCCTGGTGTGTATACCGGTCTTTTCATTTGAGAGTGAATTTGTATGCAAGCAT 1113
 Db 1163 TGTGTGGTGTGGCCAGCGCTCACTGCACTGCAAGATGAGCTGTGCTGTAGGACT 1222
 QY 1114 ATATATATGACGAGACAGATCCAAATTAAGAGGCTTCTTGAATCTAATATGCTATCC 1173
 Db 1223 TAGTCTTGATGAAGAGAGGAGGAGGAGGAGCTGTGGGGAATGGGGGTGCTGCTGC 1282
 QY 1174 TGCAGCTGTGAAGAAATACAGAGCATGAGCTGGAAGAAAGAAATTTGTCACTCTCA 1233
 Db 1283 TGCAGCTGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1342
 QY 1234 AAGCTATAGCTTGTGTAATTCAGACTCCATGCAATGAGAGAT--CTTGAAGCGCTTCA 1291
 Db 1343 AGGCTTGGGCTTGTGCAATTCAGACTGTGTGCAATGAGAGATGAGGCGGAGGCTGTGA 1402

QY 1292 GAGCTTC-AGGATGTCTTACATGAGGCGCTGCAGATTTATGAGCTGGCCAG----- 1343
 Db 1403 GCAAGCTGGAGAAAGCTCCTGACGAGGCGCTGTGAGATGAGACCGGCGGCTGGGCC 1462
 QY 1344 -----CATATGAAAGACCTCGTGCAGGTGGCAAGATGTGATGACACTGCCACTGCTGA 1398
 Db 1463 CCGAGAGGGGTGTGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1522
 QY 1399 GCGAGACTCTTACAAAGGCGCTGACGATTTCTTACAAATCAATCAAACTAGAAAGCAAGTCC 1458
 Db 1523 GCGAGACAGCGGCAAAAGTGTGCGCCATTCTATGGGTGAGAGCTGAGAGGCAAGGTGC 1582
 QY 1459 CAATGCAAAACTTTTGTGAAAATGTTGAGGCGCAAGT 1498
 Db 1583 CAATGCAAAAGCTTTTGTGAGATGTGAGGCGCATAT 1622

Search completed: August 10, 2004, 05:39:35
 Job time : 2420 secs

CC sequence represents a human nNR2 protein
 XX Sequence 458 AA;

Query Match 100.0%; Score 2388; DB 2; Length 458;
 Best Local Similarity 100.0%; Pred. No. 9.2e-217;
 Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDSVELCLPESFSLHYEELLCRMNKRHDSSCSSEFIKTEPSPASLTDSVNHSPGG 60
DB 1 MDSVELCLPESFSLHYEELLCRMNKRHDSSCSSEFIKTEPSPASLTDSVNHSPGG 60
QY 61 SSDASGYSSTMGHONGLDSPPLYPAPILGSGGPVKLYDDCSSTIVEDPOTKCEYML 120
DB 61 SSDASGYSSTMGHONGLDSPPLYPAPILGSGGPVKLYDDCSSTIVEDPOTKCEYML 120
QY 121 NSMPKRLCLVCGDIASGYHGVASCEACKAFKRTIQNIEYSCPATNECEITRRRSC 180
DB 121 NSMPKRLCLVCGDIASGYHGVASCEACKAFKRTIQNIEYSCPATNECEITRRRSC 180
QY 181 QACRFMKCLKVGMLKEGVRLDRVGRQKRYRRIDAENSPLYLNQVOPAKKPYNKIYSH 240
DB 181 QACRFMKCLKVGMLKEGVRLDRVGRQKRYRRIDAENSPLYLNQVOPAKKPYNKIYSH 240
QY 241 LVAPEPKIYAMPDPTVPDSDIKALITLCLADRELVIIGMAKHIPGFSTLSADQMSL 300
DB 241 LVAPEPKIYAMPDPTVPDSDIKALITLCLADRELVIIGMAKHIPGFSTLSADQMSL 300
QY 301 LOSAMMELLILGVYRSLSFEDLVYADDTIMEDQSKLAGLDLNNALIQLVKKYYSMK 360
DB 301 LOSAMMELLILGVYRSLSFEDLVYADDTIMEDQSKLAGLDLNNALIQLVKKYYSMK 360
QY 361 LEKEEFTLKAIALANDSMHIEDVEAVOKLQDYLHEALQDYEAGQHMEDEPRRAGKMLMT 420
DB 361 LEKEEFTLKAIALANDSMHIEDVEAVOKLQDYLHEALQDYEAGQHMEDEPRRAGKMLMT 420
QY 421 LPLLRQSTRAVOHFYNIKLEGKVPMKLFLEMLEAVY 458
DB 421 LPLLRQSTRAVOHFYNIKLEGKVPMKLFLEMLEAVY 458

```

RESULT 2

AAB09965 ID AAB09965 standard; protein; 458 AA.

XX AAB09965;

DT 19-OCT-2000 (first entry)

XX Human ERGamma protein.

XX Human; ERGamma; brain; estrogen-related receptor gamma; cytosolic;

XX lipid metabolism; fatty acid synthesis; antiarteriosclerotic; treatment;

XX drug development; diabetes.

XX Homo sapiens.

XX WO200026365-A1.

XX 11-MAY-2000.

XX 02-NOV-1999; 99WO-JP006097.

XX 04-NOV-1998; 98JP-00313194.

XX (KAZU-) KAZUSA DNA RES INST FOUND.

XX (TAIS) TAISHO PHARM CO LTD.

XX Ohara O, Nagase T, Nomura N, Takayama K, Toyoda H, Yoshimoto M;

XX WPI; 2000-365614/31.

XX DR N-FSDB; AAA40078, AAA40079.

PT An estrogen-related receptor gamma protein with lipid metabolism
 PT regulatory and fatty acid synthesis functions, and its encoding gene,
 PT useful as drugs and in developing drugs for treatment of e.g.
 PT arteriosclerosis and diabetes.

XX Claim 1a; Page 22-24; 38pp; Japanese.

CC This invention describes a novel estrogen-related receptor gamma
 CC (ERgamma) protein (I) and its variants maintaining receptor activity,
 CC having lipid metabolism regulatory and fatty acid synthesis functions.
 CC The product of the invention has cytosolic and antiarteriosclerotic
 CC activity. The protein and gene are useful as drugs and in developing
 CC drugs for treatment of e.g. arteriosclerosis and diabetes. The protein is
 CC obtained by cloning human brain-originated cDNA library. This sequence
 CC represents the human brain ERgamma protein which is described in the
 CC method of the invention

XX Sequence 458 AA;

Query Match 100.0%; Score 2388; DB 3; Length 458;
 Best Local Similarity 100.0%; Pred. No. 9.2e-217;
 Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MDSVELCLPESFSLHYEELLCRMNKRHDSSCSSEFIKTEPSPASLTDSVNHSPGG 60
DB 1 MDSVELCLPESFSLHYEELLCRMNKRHDSSCSSEFIKTEPSPASLTDSVNHSPGG 60
QY 61 SSDASGYSSTMGHONGLDSPPLYPAPILGSGGPVKLYDDCSSTIVEDPOTKCEYML 120
DB 61 SSDASGYSSTMGHONGLDSPPLYPAPILGSGGPVKLYDDCSSTIVEDPOTKCEYML 120
QY 121 NSMPKRLCLVCGDIASGYHGVASCEACKAFKRTIQNIEYSCPATNECEITRRRSC 180
DB 121 NSMPKRLCLVCGDIASGYHGVASCEACKAFKRTIQNIEYSCPATNECEITRRRSC 180
QY 181 QACRFMKCLKVGMLKEGVRLDRVGRQKRYRRIDAENSPLYLNQVOPAKKPYNKIYSH 240
DB 181 QACRFMKCLKVGMLKEGVRLDRVGRQKRYRRIDAENSPLYLNQVOPAKKPYNKIYSH 240
QY 241 LVAPEPKIYAMPDPTVPDSDIKALITLCLADRELVIIGMAKHIPGFSTLSADQMSL 300
DB 241 LVAPEPKIYAMPDPTVPDSDIKALITLCLADRELVIIGMAKHIPGFSTLSADQMSL 300
QY 301 LOSAMMELLILGVYRSLSFEDLVYADDTIMEDQSKLAGLDLNNALIQLVKKYYSMK 360
DB 301 LOSAMMELLILGVYRSLSFEDLVYADDTIMEDQSKLAGLDLNNALIQLVKKYYSMK 360
QY 361 LEKEEFTLKAIALANDSMHIEDVEAVOKLQDYLHEALQDYEAGQHMEDEPRRAGKMLMT 420
DB 361 LEKEEFTLKAIALANDSMHIEDVEAVOKLQDYLHEALQDYEAGQHMEDEPRRAGKMLMT 420
QY 421 LPLLRQSTRAVOHFYNIKLEGKVPMKLFLEMLEAVY 458
DB 421 LPLLRQSTRAVOHFYNIKLEGKVPMKLFLEMLEAVY 458

```

RESULT 3

AA93823 ID AA93823 standard; protein; 458 AA.

XX AA93823;

DT 05-JUL-2000 (first entry)

XX Human nuclear receptor nNR2.

XX Human nuclear receptor protein-2; nNR2; physiological function;

XX cell development and differentiation controller; gene expression.

XX Homo sapiens.

XX US6054295-A.

PD 25-APR-2000.
 XX 26-AUG-1998; 98US-00141000.
 XX
 PR 27-AUG-1997; 97US-0057090P.
 PR 21-OCT-1997; 97US-0062922P.
 PR 19-MAR-1998; 98US-0078633P.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Chen F;
 XX
 DR WPI: 2000-328352/28.
 DR N-PSDB; AAA09801.
 XX
 PT New polynucleotide encoding human nuclear receptor protein-1 is useful
 CC for screening compounds that acts as modulators of cell differentiation,
 CC cell development and physiological function.
 XX
 PS Disclosure; Fig 5; 58pp; English.
 XX
 CC This sequence represents the human nuclear receptor protein-2 (nR2). The
 CC protein is a cell development and differentiation controller which is
 CC useful for screening compounds that act as modulators of cell
 CC differentiation, cell development and physiological functions or for
 CC treating and diagnosing disorders associated with altered expression of
 CC the DNA
 XX
 SQ Sequence 458 AA;
 XX
 Query Match 100.0%; Score 2388; DB 3; Length 458;
 Best Local Similarity 100.0%; Pred. No. 9.2e-217; Indels 0; Gaps 0;
 Matches 458; Conservative 0; Mismatches 0;
 QY 1 MSVELCLPESLSLHHEEELCRMNKKDHIDSSCSFRTKTEPSSPASITDSVNHHSPPG 60
 DB 1 MSVELCLPESLSLHHEEELCRMNKKDHIDSSCSFRTKTEPSSPASITDSVNHHSPPG 60
 QY 61 SSDASGSYSTNMGHONGDLPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPOTKCEYML 120
 DB 61 SSDASGSYSTNMGHONGDLPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPOTKCEYML 120
 QY 121 NSMPKRLCYCGDIASGYHYGVAACACAFRRITIQNIEVSCPATNECEITKRRKSC 180
 DB 121 NSMPKRLCYCGDIASGYHYGVAACACAFRRITIQNIEVSCPATNECEITKRRKSC 180
 QY 181 QACRFKCLKVGMLKEGVRLDRVRGGRQYKRRIDAENSPYINPOLVOPAKKPYNKIVSH 240
 DB 181 QACRFKCLKVGMLKEGVRLDRVRGGRQYKRRIDAENSPYINPOLVOPAKKPYNKIVSH 240
 QY 241 LVAABPEKTYAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHLPGFSTLSLAOMSL 300
 DB 241 LVAABPEKTYAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHLPGFSTLSLAOMSL 300
 QY 301 LGSAMWEIILGVVYNSLSFEDELVIADYIMDEDSKLAGLDINNAIIQLVKKYKSMK 360
 DB 301 LGSAMWEIILGVVYNSLSFEDELVIADYIMDEDSKLAGLDINNAIIQLVKKYKSMK 360
 QY 361 LEKEEVSVTALKALANDSMHIEDVAVOQLQOVLEHALODYEAQOMHEDPRAGKMLMT 420
 DB 361 LEKEEVSVTALKALANDSMHIEDVAVOQLQOVLEHALODYEAQOMHEDPRAGKMLMT 420
 QY 421 LPLLRQTSKAVQHFYNIKLEGRVMEKLEMLEAKV 458
 DB 421 LPLLRQTSKAVQHFYNIKLEGRVMEKLEMLEAKV 458
 RESULT 4
 ID ADC23486 standard; protein; 458 AA.
 AC
 XX
 XX ADC23486;
 XX

DT 18-DEC-2003 (first entry)
 XX
 DE Ligand binding domain of the oestrogen related receptor 3 (ERR3) protein.
 XX
 KW oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;
 KW ligand binding domain; fertility; birth control; bone remodelling;
 KW cancer; protein coordinate data.
 XX
 OS Unidentified.
 XX
 PN WO2003064468-A2.
 XX
 PD 07-AUG-2003.
 XX
 PF 30-JAN-2003; 2003WO-BP000959.
 XX
 PR 31-JAN-2002; 2002US-0352551P.
 XX
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Moras D, Renaud J, Greschik H, Murtz J;
 XX
 DR WPI: 2003-663467/62.
 DR N-PSDB; ADC23486.
 XX
 PT New peptide fragment, useful for screening compounds that are agonists or
 CC antagonists of the transcriptional-activating activity of the estrogen-
 CC related receptor 3 (ERR3).
 XX
 PS Claim 1; SEQ ID NO 1; 259pp; English.
 XX
 CC This invention relates to novel peptide fragments that have an agonistic
 CC or antagonistic effect on the transcriptional-activating activity of the
 CC oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen
 CC related receptor gamma protein. Specifically, the peptide fragment of the
 CC invention comprises a ligand binding domain of ERR3, which becomes
 CC functionally active when fused to a protein containing a DNA binding
 CC domain and mimics the transcriptional-activating activity of the complete
 CC ERR3 protein. As such, the peptide fragment can be used in a screening
 CC method or to design and select compounds that affect ERR3 activity.
 CC Furthermore, the agonists and antagonists of ERR3 are biologically active
 CC compounds that can be used to modulate the oestrogenic response on
 CC fertility, birth control, bone remodelling, breast and prostate cancer.
 CC This polypeptide sequence is the ligand binding pocket of the ERR3
 CC protein of the invention.
 XX
 SQ Sequence 458 AA;
 XX
 Query Match 100.0%; Score 2388; DB 7; Length 458;
 Best Local Similarity 100.0%; Pred. No. 9.2e-217; Indels 0; Gaps 0;
 Matches 458; Conservative 0; Mismatches 0;
 QY 1 MSVELCLPESLSLHHEEELCRMNKKDHIDSSCSFRTKTEPSSPASITDSVNHHSPPG 60
 DB 1 MSVELCLPESLSLHHEEELCRMNKKDHIDSSCSFRTKTEPSSPASITDSVNHHSPPG 60
 QY 61 SSDASGSYSTNMGHONGDLPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPOTKCEYML 120
 DB 61 SSDASGSYSTNMGHONGDLPPLYPSPAPILGSGSPVRKLYDDCSSTIVEDPOTKCEYML 120
 QY 121 NSMPKRLCYCGDIASGYHYGVAACACAFRRITIQNIEVSCPATNECEITKRRKSC 180
 DB 121 NSMPKRLCYCGDIASGYHYGVAACACAFRRITIQNIEVSCPATNECEITKRRKSC 180
 QY 181 QACRFKCLKVGMLKEGVRLDRVRGGRQYKRRIDAENSPYINPOLVOPAKKPYNKIVSH 240
 DB 181 QACRFKCLKVGMLKEGVRLDRVRGGRQYKRRIDAENSPYINPOLVOPAKKPYNKIVSH 240
 QY 241 LVAABPEKTYAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHLPGFSTLSLAOMSL 300
 DB 241 LVAABPEKTYAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHLPGFSTLSLAOMSL 300
 QY 301 LGSAMWEIILGVVYNSLSFEDELVIADYIMDEDSKLAGLDINNAIIQLVKKYKSMK 360
 DB 301 LGSAMWEIILGVVYNSLSFEDELVIADYIMDEDSKLAGLDINNAIIQLVKKYKSMK 360

Db		301	IOSAMHEILILGVIRSLSPFEDELVADYIDNDEDSKLAGLIDNNALLIQVKTKSKM	360
Qy		361	LEKEEFVTIKALANSDSMHIEDVEAVOKLODVHEALOQYEAGOMEDPRRAGOLMT	420
Db		361	LEKEEFVTIKALANSDSMHIEDVEAVOKLODVHEALOQYFAGOHMEDPRRAGOLMT	420
Qy		421	LPLAQSTSKVAOHFNITLKGVPNMKKFLMLEAKV	458
Db		421	LPLRQTSTKVAOHEFYNIKLGKVPMKKFLMLEAKV	458
 RESULT 5 ADD69757				
ID	ADD69757	standard; protein; 458 AA.		
XX	ADD69757;			
AC				
XX				
DT	15-JAN-2004	(first entry)		
XX				
DE		Human ERR gamma 3-related protein - SEQ ID 6.		
XX				
KW		nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;		
KW		oestrogen receptor; ER; thyroid hormone; TR; human.		
OS		Homo sapiens.		
XX				
PN	WO2003080831-AI.			
PD				
XX				
XX	02-OCT-2003.			
PE				
XX	25-MAR-2003; 2003WO-JP003611.			
FR				
XX	25-MAR-2002; 2002JP-00084560.			
PA				
XX	(FUJII) FUJISAWA PHARM CO LTD.			
PI				
DR	Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;			
DR	WPI; 2003-779262/73.			
DR	N-PSDB; ADD69756.			
PT		Polynucleotides encoding nuclear receptors, and the encoded proteins,		
PT		useful as diagnostic agents, and for identification of agents that affect		
XX		receptor activity.		
PS				
XX	Claim 9; SEQ ID NO 6; 148pp; Japanese.			
CC		The invention relates to novel nuclear receptor ERR (oestrogen receptor-		
CC		related receptor) gamma 3 polynucleotides. The polynucleotides of the		
CC		invention may be useful for diagnosis of disorders caused by abnormal		
CC		nuclear receptor activity, particularly those related to abnormal		
CC		oestrogen receptor (ER), ERR or thyroid hormone receptor (TR) activity.		
CC		Furthermore, the polynucleotides and proteins may be useful for		
CC		evaluating agents that affect the activity of nuclear receptors. The		
CC		current sequence is that of the human ERR gamma 3-related protein of the		
XX		invention.		
XX				
XX	Sequence 458 AA;			
Query March	100.0%; Score 2388; DB 7; Length 458;			
Best Local Similarity	100.0%; Pred. No. 9.2e-217;			
Matches 458; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MDSVELCLPESFSLSHYEEELLCRMENKORHIDSSCSFKTPESSPASLTDSVNHHSPG	60	
Db	1	MDSVELCLPESFSLSHYEEELLCRMENKORHIDSSCSFKTPESSPASLTDSVNHHSPG	60	
Qy	61	SSDASGSYSSTMNGHONGLDSPPLYPAPILIGSGPVKKLYDDCSSTIVEDPOTCEYWL	120	
Db	61	SSDASGSYSSTMNHQNGLDSPPLYPAPILIGSGFPVKKLYDDCSSTIVEDPOTCEYWL	120	
Qy	121	NSMPKRCLCIVCGDIASGYHVYASCEACKAFPKRTIQGNIEYSCPATNECETTKRRRSKC	180	

Db	121	NSMPKRLCLVCDIDIASGHHYGVASCEACKAFKKRTIQGNIIEYSCPATNECEITKRRRKS	180
Qy	181	QACRFMKCLKVMLEKGVLLDRVGRGRKKRIRIDENSPLYNPOLVOPAKRPYNTVSH	240
Db	181	QACRFMKCLKVMLEKGVLLDRVGRGRKKRIRIDENSPLYNPOLVOPAKRPYNTVSH	240
Qy	241	ILVAPEPKIYAMPDETPVDSIDIKALTLLCDLADRELHVLIIGNAKHIPGSTSLADQMSL	300
Db	241	ILVAPEPKIYAMPDETPVDSIDIKALTLLCDLADRELHVLIIGNAKHIPGSTSLADQMSL	300
Qy	301	LOSAMELIIILGVVRSLSFEDELYADDIYIMDEDSKLAGLIDNNAILQLVKRYKSK	360
Db	301	LOSAMELIIILGVVRSLSFEDELYADDIYIMDEDSKLAGLIDNNAILQLVKRYKSK	360
Qy	361	LEKEFEVTLKATALNSDSMHIIEYAVOKLQDVHEALQDYEAQGMEDPPRAGCLMT	420
Db	361	LEKEFEVTLKATALNSDSMHIIEYAVOKLQDVHEALQDYEAQGMEDPPRAGCLMT	420
Qy	421	LPILRQSTKAVQHFNYIKLBSKVPMHKLFLMLEAKV	458
Db	421	LPILRQSTKAVQHFNYIKLBSKVPMHKLFLMLEAKV	458
RESULT 6			
ID	ADCC23494	standard; protein; 458 AA.	
XX	AC	ADCC23494;	
XX	DT	18-DEC-2003 (first entry)	
DE	XX	Mutant ligand binding domain of the ERR3 protein - A272L.	
XX	XX	oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;	
KW	XX	ligand binding domain; fertility; birth control; bone remodelling;	
KM	XX	cancer; protein coordinate data; mutant; mutein.	
XX	OS	Unidentified.	
FE	FE	Key Location/Qualifiers	
FT	FT	Misc-difference 272	
XX	XX	/note= "wild type Ala substituted for Leu"	
PN	XX	WO2003064468-A2.	
PD	XX	07-AUG-2003.	
PF	XX	30-JAN-2003; 2003WO-EP000959.	
XX	XX	31-JAN-2002; 2002US-0352551P.	
XX	PA	(CNRS) CNRS CENT NAT RECH SCI.	
PI	XX	Moras D, Renaud J, Greschik H, Wurtz J;	
DR	XX	WPI; 2003-663467/62.	
PT	XX	New peptide fragment, useful for screening compounds that are agonists or	
PT	XX	antagonists of the transcriptional-activating activity of the estrogen-	
XX	XX	related receptor 3 (ERR3).	
XX	XX	Claim 49; Page; 259pp; English.	
XX	XX	This invention relates to novel peptide fragments that have an agonistic	
CC	XX	or antagonistic effect on the transcriptional-activating activity of the	
CC	XX	oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen	
CC	XX	related receptor gamma protein. Specifically, the peptide fragment of the	
CC	XX	invention comprises a ligand binding domain of ERR3, which becomes	
CC	XX	functionally active when fused to a protein containing a DNA binding	
CC	XX	domain and mimics the transcriptional-activating activity of the complete	
CC	XX	ERR3 protein. As such, the peptide fragment can be used in a screening	
CC	XX	method or to design and select compounds that affect ERR3 activity.	

CC Furthermore, the agonists and antagonists of ERR3 are biologically active
CC compounds that can be used to modulate the oestrogenic response on
CC fertility, birth control, bone remodeling, breast and prostate cancer.
CC This polypeptide sequence is a mutant protein (A272J) comprising the
CC ligand binding pocket of the ERR3 protein of the invention. NOTE: This
CC sequence is not given in the specification but is derived from
CC information given in the claim 49.

XX Sequence 458 AA;

Query Match 99.8%; Score 2383; DB 7; Length 458;
Best Local Similarity 99.8%; Pred. No. 2.7e-216;
Matches 457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEBELLCRMNKNDRHIDSSCSFTKTEPSPASLTDSVNHHSFG 60
Db 1 MDSVELCLPESFSLHYEBELLCRMNKNDRHIDSSCSFTKTEPSPASLTDSVNHHSFG 60
QY 61 SSDASGYSSTMGHONGLDSPPLYPSPAPILGSGSPVAKLYDDCSSTIVEDPQTKCEYML 120
Db 61 SSDASGYSSTMGHONGLDSPPLYPSPAPILGSGSPVAKLYDDCSSTIVEDPQTKCEYML 120
QY 121 NSMFKRLCLVCGDIASGYHYGVAACEAKAFKRTIQGNIEYSCPATNECEITRRRRKSC 180
Db 121 NSMFKRLCLVCGDIASGYHYGVAACEAKAFKRTIQGNIEYSCPATNECEITRRRRKSC 180
QY 181 QACRFMKCLKYGMLKEGYRLDRVRGROKYRRIDAENSPLYNPOLVOPAKKPYNKIVSH 240
Db 181 QACRFMKCLKYGMLKEGYRLDRVRGROKYRRIDAENSPLYNPOLVOPAKKPYNKIVSH 240
QY 241 LLVAPEKTIYAMPPTVPDSDIKALTTCLADRELAVIIGWAKHIFPSTLSLADQMSL 300
Db 241 LLVAPEKTIYAMPPTVPDSDIKALTTCLADRELAVIIGWAKHIFPSTLSLADQMSL 300
QY 301 LOSAMELILIGVYRSLSPFEDELVADYIMDEDSQSLAGLDLNNAILQLVKKYKSMK 360
Db 301 LOSAMELILIGVYRSLSPFEDELVADYIMDEDSQSLAGLDLNNAILQLVKKYKSMK 360
QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLDVHEALQDYEAQHMEDPRRACKMLMT 420
Db 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLDVHEALQDYEAQHMEDPRRACKMLMT 420
QY 421 LPILRQSTKAVQHFYNIKLGKVPMMKLFLEMLAKV 458
Db 421 LPILRQSTKAVQHFYNIKLGKVPMMKLFLEMLAKV 458

RESULT 7
ADC23493
ID ADC23493 standard; protein; 458 AA.

XX ADC23493;
XX 18-DEC-2003 (first entry)
XX Mutant ligand binding domain of the ERR3 protein - A272F.
XX oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;
XX ligand binding domain; fertility; birth control; bone remodeling;
XX cancer; protein coordinate data; mutant; mutein.
XX Unidentified.
XX Key Location/Qualifiers
XX Misc-difference 272 /note="Wild type Ala substituted for Phe"
XX WO2003064468-A2.
XX 07-AUG-2003.
XX 30-JAN-2003; 2003WO-EP000959.

PR 31-JAN-2002; 2002US-0352551P.
XX (CNRS) CNRS CENT NAT RECH SCT.
XX Moras D, Renaud J, Greshik H, Wurtz J;
XX WPI; 2003-663467/62.
XX New peptide fragment, useful for screening compounds that are agonists or
XX antagonists of the transcriptional-activating activity of the estrogen-
XX related receptor 3 (ERR3).
PS Claim 49; Page; 259pp; English.

CC This invention relates to novel peptide fragments that have an agonistic
CC or antagonistic effect on the transcriptional-activating activity of the
CC oestrogen related receptor 3 (ERR3) protein, also known as the oestrogen
CC related receptor gamma protein. Specifically, the peptide fragment of the
CC invention comprises a ligand binding domain of ERR3, which becomes
CC functionally active when fused to a protein containing a DNA binding
CC domain and mimics the transcriptional-activating activity of the complete
CC ERR3 protein. As such, the peptide fragment can be used in a screening
CC method or to design and select compounds that affect ERR3 activity.
CC Furthermore, the agonists and antagonists of ERR3 are biologically active
CC compounds that can be used to modulate the oestrogenic response on
CC fertility, birth control, bone remodeling, breast and prostate cancer.
CC This polypeptide sequence is a mutant protein (A272F) comprising the
CC ligand binding pocket of the ERR3 protein of the invention. NOTE: This
CC sequence is not given in the specification but is derived from
CC information given in the claim 49.

XX Sequence 458 AA;

Query Match 99.7%; Score 2382; DB 7; Length 458;
Best Local Similarity 99.8%; Pred. No. 3.4e-216;
Matches 457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEBELLCRMNKNDRHIDSSCSFTKTEPSPASLTDSVNHHSFG 60
Db 1 MDSVELCLPESFSLHYEBELLCRMNKNDRHIDSSCSFTKTEPSPASLTDSVNHHSFG 60
QY 61 SSDASGYSSTMGHONGLDSPPLYPSPAPILGSGSPVAKLYDDCSSTIVEDPQTKCEYML 120
Db 61 SSDASGYSSTMGHONGLDSPPLYPSPAPILGSGSPVAKLYDDCSSTIVEDPQTKCEYML 120
QY 121 NSMFKRLCLVCGDIASGYHYGVAACEAKAFKRTIQGNIEYSCPATNECEITRRRRKSC 180
Db 121 NSMFKRLCLVCGDIASGYHYGVAACEAKAFKRTIQGNIEYSCPATNECEITRRRRKSC 180
QY 181 QACRFMKCLKYGMLKEGYRLDRVRGROKYRRIDAENSPLYNPOLVOPAKKPYNKIVSH 240
Db 181 QACRFMKCLKYGMLKEGYRLDRVRGROKYRRIDAENSPLYNPOLVOPAKKPYNKIVSH 240
QY 241 LLVAPEKTIYAMPPTVPDSDIKALTTCLADRELAVIIGWAKHIFPSTLSLADQMSL 300
Db 241 LLVAPEKTIYAMPPTVPDSDIKALTTCLADRELAVIIGWAKHIFPSTLSLADQMSL 300
QY 301 LOSAMELILIGVYRSLSPFEDELVADYIMDEDSQSLAGLDLNNAILQLVKKYKSMK 360
Db 301 LOSAMELILIGVYRSLSPFEDELVADYIMDEDSQSLAGLDLNNAILQLVKKYKSMK 360
QY 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLDVHEALQDYEAQHMEDPRRACKMLMT 420
Db 361 LEKEEFVTLKAIALANSDSMHIEDVEAVQKLDVHEALQDYEAQHMEDPRRACKMLMT 420
QY 421 LPILRQSTKAVQHFYNIKLGKVPMMKLFLEMLAKV 458
Db 421 LPILRQSTKAVQHFYNIKLGKVPMMKLFLEMLAKV 458

RESULT 8
ADC23492
ID ADC23492 standard; protein; 458 AA.

```

XX AC ADC23492;
XX DT 18-DEC-2003 (first entry)
XX DE Mutant ligand binding domain of the ERR3 protein - F435L.
XX KM oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;
XX KM ligand binding domain; fertility; birth control; bone remodeling;
XX KM cancer; protein coordinate data; mutant; mutein.
XX OS Unidentified.
XX PS Key Location/Qualifiers
XX FT MISC-difference 435 /note= "Wild type Phe substituted for Leu"
XX PN W02003064468-A2.
XX PD 07-AUG-2003.
XX PE 30-JAN-2003; 2003WO-EP000959.
XX PR 31-JAN-2002; 2002US-0352551P.
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX PI Moras D, Renaud J, Greschik H, Wurtz J;
XX DR WPI; 2003-663467/62.
XX PT New peptide fragment, useful for screening compounds that are agonists or
XX PT antagonists of the transcriptional-activating activity of the estrogen-
XX PT related receptor 3 (ERR3).
XX PS Claim 49; Page; 259pp; English.
XX CC This invention relates to novel peptide fragments that have an agonistic
XX CC or antagonistic effect on the transcriptional-activating activity of the
XX CC estrogen related receptor 3 (ERR3) protein, also known as the oestrogen
XX CC related receptor gamma protein. Specifically, the peptide fragment of the
XX CC invention comprises a ligand binding domain of ERR3, which becomes
XX CC functionally active when fused to a protein containing a DNA binding
XX CC domain and mimics the transcriptional-activating activity of the complete
XX CC ERR3 protein. As such, the peptide fragment can be used in a screening
XX CC method or to design and select compounds that affect ERR3 activity.
XX CC Furthermore, the agonists and antagonists of ERR3 are biologically active
XX CC compounds that can be used to modulate the oestrogenic response on
XX CC fertility, birth control, bone remodeling, breast and prostate cancer.
XX CC This polypeptide sequence is a mutant protein (F435L) comprising the
XX CC ligand binding pocket of the ERR3 protein of the invention. NOTE: This
XX CC sequence is not given in the specification but is derived from
XX CC information given in the claim 49.
XX SQ Sequence 458 AA;
XX
XX Query Match 99.7%; Score 2382; DB 7; Length 458;
XX Best Local Similarity 99.8%; Pred. No. 3.4e-216;
XX Matches 457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MDSVELCLPESFSLHYEEELLCRMSNDRHIDSSCSFFIKTEPSSPASTLDSVNHSPGG 60
XX DB 1 MDSVELCLPESFSLHYEEELLCRMSNDRHIDSSCSFFIKTEPSSPASTLDSVNHSPGG 60
XX QY 61 SPSASGSYSTMGCHQGLDSPPIYPSAPILGGSGPRKLYDDCSSITVADPQTKCYML 120
XX DB 61 SPSASGSYSTMGCHQGLDSPPIYPSAPILGGSGPRKLYDDCSSITVADPQTKCYML 120
XX QY 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFPKRTIOGNIYESCPATNECEITKRRKSC 180
XX DB 121 NSMPKRLCLVCGDIASGYHYGVAACEACKAFPKRTIOGNIYESCPATNECEITKRRKSC 180
XX QY 181 QACRFMKCLKVGMLEKEGVRLDRVYRGQRQYKRRIDAEKSPYLNQVLQVPAKKPYNKLVSH 240

```

```

XX DB 181 QACRFMKCLKVGMLEKEGVRLDRVYRGQRQYKRRIDAEKSPYLNQVLQVPAKKPYNKLVSH 240
XX QY 241 LVAEPKTIYAMPDPYVPSDDIKALTTLCDLADRELVIYTGMAKHITGFSTLSADQMSL 300
XX DB 241 LVAEPKTIYAMPDPYVPSDDIKALTTLCDLADRELVIYTGMAKHITGFSTLSADQMSL 300
XX QY 301 LOSAMMEILLGVVYRSLSEDELVVADYIMDEDSKLAGLDNNNAIILQVKKYSMK 360
XX DB 301 LOSAMMEILLGVVYRSLSEDELVVADYIMDEDSKLAGLDNNNAIILQVKKYSMK 360
XX QY 361 LEKEEFVTLKALALANDSMHIEDVEAVQKLDVLEHALQDYEAQGMEDEPRRAGKMLMT 420
XX DB 361 LEKEEFVTLKALALANDSMHIEDVEAVQKLDVLEHALQDYEAQGMEDEPRRAGKMLMT 420
XX QY 421 LPLIRQSTKAVQHFNYIKLEKGVPMKLTLEMLEAVY 458
XX DB 421 LPLIRQSTKAVQHFNYIKLEKGVPMKLTLEMLEAVY 458
XX
XX RESULT 9
XX ID ADC23495 standard; protein; 458 AA.
XX AC ADC23495;
XX DT 18-DEC-2003 (first entry)
XX DE Mutant ligand binding domain of the ERR3 protein - I345I/ F435L/ F450L.
XX KM oestrogen related receptor 3; ERR3; oestrogen related receptor gamma;
XX KM ligand binding domain; fertility; birth control; bone remodeling;
XX KM cancer; protein coordinate data; mutant; mutein.
XX OS Unidentified.
XX PS Key Location/Qualifiers
XX FT MISC-difference 345 /note= "Wild type Leu substituted for Ile"
XX FT MISC-difference 435 /note= "Wild type Phe substituted for Leu"
XX FT MISC-difference 450 /note= "Wild type Phe substituted for Leu"
XX PN W02003064468-A2.
XX PD 07-AUG-2003.
XX PE 30-JAN-2003; 2003WO-EP000959.
XX PR 31-JAN-2002; 2002US-0352551P.
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX PI Moras D, Renaud J, Greschik H, Wurtz J;
XX DR WPI; 2003-663467/62.
XX PT New peptide fragment, useful for screening compounds that are agonists or
XX PT antagonists of the transcriptional-activating activity of the estrogen-
XX PT related receptor 3 (ERR3).
XX PS Claim 49; Page; 259pp; English.
XX CC This invention relates to novel peptide fragments that have an agonistic
XX CC or antagonistic effect on the transcriptional-activating activity of the
XX CC estrogen related receptor 3 (ERR3) protein, also known as the oestrogen
XX CC related receptor gamma protein. Specifically, the peptide fragment of the
XX CC invention comprises a ligand binding domain of ERR3, which becomes
XX CC functionally active when fused to a protein containing a DNA binding
XX CC domain and mimics the transcriptional-activating activity of the complete
XX CC ERR3 protein. As such, the peptide fragment can be used in a screening
XX CC method or to design and select compounds that affect ERR3 activity.

```

CC Furthermore, the agonists and antagonists of ERR3 are biologically active
 CC compounds that can be used to modulate the oestrogenic response on
 CC fertility, birth control, bone remodeling, breast and prostate cancer.
 CC This polypeptide sequence is a mutant protein (I345I/ F450L)
 CC comprising the ligand binding pocket of the ERR3 protein of the
 CC invention. NOTE: This sequence is not given in the specification but is
 CC derived from information given in the claim 49.

XX Sequence 458 AA;

Query Match 99.4%; Score 2374; DB 7; Length 458;
 Best Local Similarity 99.3%; Pred. No. 1.9e-215;
 Matches 455; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSVELCLPESRSLHYEEELCRMSNKRHIDS CSSPFIKTEPSSPASTLDSVNHSFGG 60
 DB 1 MDSVELCLPESRSLHYEEELCRMSNKRHIDS CSSPFIKTEPSSPASTLDSVNHSFGG 60
 QY 61 SSDASGYSSTMNGHONGIDSPPLYPSPAPILGSGPVRKLYDGSSTIVDPQTCETML 120
 DB 61 SSDASGYSSTMNGHONGIDSPPLYPSPAPILGSGPVRKLYDGSSTIVDPQTCETML 120
 QY 121 NSMPRLCLVCGDIASGYHGVASCEACAPFKRTIOGNIYSCPATNECETTKRRRSC 180
 DB 121 NSMPRLCLVCGDIASGYHGVASCEACAPFKRTIOGNIYSCPATNECETTKRRRSC 180
 QY 181 QACRFMKCLXVGMLEKGVRLDRVGRGROKYKRRIDAENSPYLPOLVOPAKKPYNKIYSH 240
 DB 181 QACRFMKCLXVGMLEKGVRLDRVGRGROKYKRRIDAENSPYLPOLVOPAKKPYNKIYSH 240
 QY 241 LLVAPPEKTIYAMPDPTVPDSIDKALTTLCDLADRELVIITIGAKHITPGSTISLADQSL 300
 DB 241 LLVAPPEKTIYAMPDPTVPDSIDKALTTLCDLADRELVIITIGAKHITPGSTISLADQSL 300
 QY 301 LOSAMEIILIGVYRSLSFEDELVYADYIMDEPOSKLAGLIDNNAILQVKKYKSMK 360
 DB 301 LOSAMEIILIGVYRSLSFEDELVYADYIMDEPOSKLAGLIDNNAILQVKKYKSMK 360
 QY 361 LEKEEFVTLKATALANSDSMHIEDVEAVOKLQDVLHEALQDYEAGQHMDPPRAGKMLMT 420
 DB 361 LEKEEFVTLKATALANSDSMHIEDVEAVOKLQDVLHEALQDYEAGQHMDPPRAGKMLMT 420
 QY 421 LPLAQTSITKAVQHFNITKLEKVPMEHKLFLMLEAKV 458
 DB 421 LPLAQTSITKAVQHFNITKLEKVPMEHKLFLMLEAKV 458

RESULT 10

AA661192 standard; protein; 435 AA.

XX AA661192;

DT 07-DEC-1998 (first entry)

DE Steroid hormone receptor homologue HE8AB36.

XX HE8AN36; steroid hormone receptor; human; transcription factor;
 XX inflammation; arthritis; autoimmune disease; diabetes;
 KW transplant rejection; graft versus host disease; cancer;
 KW reproductive disorder; obesity; atherosclerosis; gyrate atrophy; therapy;
 KW diagnosis.

XX Homo sapiens.

XX EP866127-A2.

XX 23-SEP-1998.

XX 13-MAR-1998; 98EP-00301914.

XX 17-MAR-1997; 97GB-00005451.

PR 16-FEB-1998; 98GB-00003289.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX Mathias SL;

XX WPI: 1998-482963/42.

DR N-FSDB; AAV47645.

PT New isolated polypeptide(s) are steroid hormone receptor homologues -
 PT used for treating inflammation, arthritis, auto-immune disease, diabetes,
 PT transplant rejection etc.

PS Claim 4; Page 12; 20pp; English.

CC This polypeptide comprises human HE8AB36, a member of the nuclear hormone
 CC receptor family of polypeptides that are ligand regulated transcription
 CC factors involved in the regulation of cellular homeostasis and
 CC differentiation via the modulation of gene transcription. Its amino acid
 CC sequence was deduced from an isolated cDNA clone (see AAV47645). HE8AN36
 CC polypeptides and polymucleotides may be used in the treatment of chronic
 CC and acute inflammation, arthritis, autoimmune diseases, diabetes,
 CC transplant rejection, graft versus host disease, reproductive disorders,
 CC cancer, obesity, atherosclerosis, gyrate atrophy and other visual
 CC disorders. The invention also relates to methods for identifying agonists
 CC and antagonists/inhibitors, and for treating conditions associated with
 CC HE8AN36 imbalance using such compounds. Diagnostic assays for detecting
 CC diseases associated with inappropriate HE8AN36 activity or levels are
 CC also provided

SQ Sequence 435 AA;

Query Match 94.8%; Score 2265; DB 2; Length 435;
 Best Local Similarity 100.0%; Pred. No. 3.7e-205;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MSNKRHIDS CSSPFIKTEPSSPASTLDSVNHSFGSGSDASGYSSTMNGHONGIDSP 83
 DB 1 MSNKRHIDS CSSPFIKTEPSSPASTLDSVNHSFGSGSDASGYSSTMNGHONGIDSP 60
 QY 84 LYPSPAPILGSGPVRKLYDGSSTIVDPQTCETMLNSMPRLCLVCGDIASGYHGYA 143
 DB 61 LYPSPAPILGSGPVRKLYDGSSTIVDPQTCETMLNSMPRLCLVCGDIASGYHGYA 120
 QY 144 SCEACAPFKRTIOGNIYSCPATNECETTKRRRSCQACRFMKCLXVGMLEKGVRLDRV 203
 DB 121 SCEACAPFKRTIOGNIYSCPATNECETTKRRRSCQACRFMKCLXVGMLEKGVRLDRV 180
 QY 204 RGRGROKYKRRIDAENSPYLPOLVOPAKKPYNKIYSHLLVAPPEKTIYAMPDPTVPDSIDK 263
 DB 181 RGRGROKYKRRIDAENSPYLPOLVOPAKKPYNKIYSHLLVAPPEKTIYAMPDPTVPDSIDK 240
 QY 264 ALLTLCIDLADRELVIITIGAKHITPGSTISLADQSLLOSAMEIILIGVYRSLSFED 323
 DB 241 ALLTLCIDLADRELVIITIGAKHITPGSTISLADQSLLOSAMEIILIGVYRSLSFED 300
 QY 324 LYYADYIMDEPOSKLAGLIDNNAILQVKKYKSMKLEKEEFVTLKATALANSDSMHIE 383
 DB 301 LYYADYIMDEPOSKLAGLIDNNAILQVKKYKSMKLEKEEFVTLKATALANSDSMHIE 360
 QY 384 DVEAVOKLQDVLHEALQDYEAGQHMDPPRAGKMLMTLPLAQTSITKAVQHFNITKLEK 443
 DB 361 DVEAVOKLQDVLHEALQDYEAGQHMDPPRAGKMLMTLPLAQTSITKAVQHFNITKLEK 420
 QY 444 VPMHKLFLMLEAKV 458
 DB 421 VPMHKLFLMLEAKV 435

RESULT 11

AA82263 standard; protein; 435 AA.

XX AA82263;

XX 14-JUN-2000 (first entry)
XX Human oestrogen related receptor 3 SEQ ID NO:1.
DE Human; ligand-combining; oestrogen related receptor; ERK3; diagnosis;
KW inflammation; cancer; osteoporosis; diabetes; renal disease.
XX Homo sapiens.
XX JP2000041681-A.
XX 15-FEB-2000.
XX 31-JUL-1998; 98JP-00217933.
XX 31-JUL-1998; 98JP-00217933.
XX PR 31-JUL-1998; 98JP-00217933.
XX PA (KYOW) KYOWA HAKKO KOGYO KK.
XX WPI; 2000-295782/26.
XX DR N-PSDB; AAZ95741.
XX A new protein comprising a 435 amino acid sequence.
XX Claim 1; Page 18-19; 22pp; Japanese.
XX The present sequence represents a human oestrogen related receptor 3
CC (ERR3) protein which has ligand-combining activity. The polynucleotide
CC sequence encoding ERR3, and fragments of it, can be used for the
CC diagnosis and the treatment of inflammations, cancers, osteoporosis,
CC diabetes or renal diseases
XX Sequence 435 AA;

Query Match 94.8%; Score 2265; DB 3; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.7e-205;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 24 MSNDRHIDSSCSFITEPSSPASLTDVSVNHSPGSSDASGYSSTMGHONGLSDSP 83
DB 1 MSNDRHIDSSCSFITEPSSPASLTDVSVNHSPGSSDASGYSSTMGHONGLSDSP 60
CY 84 LYSAPILIGSGSPVKLYDDCSSTIVDPOTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143
DB 61 LYSAPILIGSGSPVKLYDDCSSTIVDPOTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 120
CY 144 SCECKAFPRKTIQGNIEYSCPATNECEITRRRRKSCQACRFMKCLKVGMLKEGVRLDRV 203
DB 121 SCECKAFPRKTIQGNIEYSCPATNECEITRRRRKSCQACRFMKCLKVGMLKEGVRLDRV 180
CY 204 RGGROKTKRRIDANSFYLNPOLOVPAKKPYNKIVSHLVAEPKITYAMPDPVPSDDIK 263
DB 181 RGGROKTKRRIDANSFYLNPOLOVPAKKPYNKIVSHLVAEPKITYAMPDPVPSDDIK 240
CY 264 ALTTLCDLADRELVIIGMAKHIPGFSTLSLADMSLQSAWMEIILIGVYRSLSEDE 323
DB 241 ALTTLCDLADRELVIIGMAKHIPGFSTLSLADMSLQSAWMEIILIGVYRSLSEDE 300
CY 324 LYYADYIMEDOSKLAGLLDNNAILQVKKYSMLKEKEEFTYTKAIALANDSMHIE 383
DB 301 LYYADYIMEDOSKLAGLLDNNAILQVKKYSMLKEKEEFTYTKAIALANDSMHIE 360
CY 384 DVEAVQKLOVLAHALQDYEAGQMEDPRRAGKMLMTLPILRQSTKAVOHFYNIKLEGK 443
DB 361 DVEAVQKLOVLAHALQDYEAGQMEDPRRAGKMLMTLPILRQSTKAVOHFYNIKLEGK 420
CY 444 VPMKILFLEMLEAKV 458
DB 421 VPMKILFLEMLEAKV 435

RESULT 12

ADD69755
ID ADD69755 standard; protein; 436 AA.
XX AC ADD69755;
XX 15-JAN-2004 (first entry)
XX Human ERR gamma 3-related protein - SEQ ID 4.
DE Human ERR gamma 3-related protein - SEQ ID 4.
KW nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;
KW oestrogen receptor; ER; thyroid hormone; TR; human.
XX Homo sapiens.
XX WO2003080831-A1.
XX 02-OCT-2003.
XX 25-MAR-2003; 2003WO-JP003611.
XX 25-MAR-2002; 2002JP-00084560.
XX (FUJII) FUJISAWA PHARM CO LTD.
XX Kojo H, Tajima K, Fukagawa M, Nishimura S, Isogai T;
XX WPI; 2003-779262/73.
XX DR N-PSDB; ADD69754.
XX Polynucleotides encoding nuclear receptors, and the encoded proteins,
PT useful as diagnostic agents, and for identification of agents that affect
PT receptor activity.
XX Claim 9; SEQ ID NO 4; 148pp; Japanese.
XX The invention relates to novel nuclear receptor ERR (oestrogen receptor-
CC related receptor) gamma 3 polynucleotides. The polynucleotides of the
CC invention may be useful for diagnosis of disorders caused by abnormal
CC nuclear receptor activity, particularly those related to abnormal
CC oestrogen receptor (ER), ERK or thyroid hormone receptor (TR) activity.
CC Furthermore, the polynucleotides and proteins may be useful for
CC evaluating agents that affect the activity of nuclear receptors. The
CC current sequence is that of the human ERR gamma 3-related protein of the
XX invention.
XX Sequence 436 AA;

Query Match 93.6%; Score 2235; DB 7; Length 436;
Best Local Similarity 98.9%; Pred. No. 2.6e-202;
Matches 430; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

XX 24 MSNDRHIDSSCSFITEPSSPASLTDVSVNHSPGSSDASGYSSTMGHONGLSDSP 83
DB 1 MSNDRHIDSSCSFITEPSSPASLTDVSVNHSPGSSDASGYSSTMGHONGLSDSP 60
CY 84 LYSAPILIGSGSPVKLYDDCSSTIVDPOTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 143
DB 61 LYSAPILIGSGSPVKLYDDCSSTIVDPOTKCEYMLNSMPKRLCLVCGDIASGYHYGVA 120
CY 144 SCECKAFPRKTIQGNIEYSCPATNECEITRRRRKSCQACRFMKCLKVGMLKEGVRLDRV 203
DB 121 SCECKAFPRKTIQGNIEYSCPATNECEITRRRRKSCQACRFMKCLKVGMLKEGVRLDRV 180
CY 204 RGGROKTKRRIDANSFYLNPOLOVPAKKPYNKIVSHLVAEPKITYAMPDPVPSDDIK 263
DB 181 RGGROKTKRRIDANSFYLNPOLOVPAKKPYNKIVSHLVAEPKITYAMPDPVPSDDIK 240
CY 264 ALTTLCDLADRELVIIGMAKHIPGFSTLSLADMSLQSAWMEIILIGVYRSLSEDE 323
DB 241 ALTTLCDLADRELVIIGMAKHIPGFSTLSLADMSLQSAWMEIILIGVYRSLSEDE 300
CY 324 LYYADYIMEDOSKLAGLLDNNAILQVKKYSMLKEKEEFTYTKAIALANDSMHIE 383

Db 301 LVVADYIMDEDSQSLAGLIDLNNAIIQLVKKYKSMKEKEFVTLKAIALANDSMHIE 360

QY 384 DVEAVOKLQDYLHEALQDYEAQGHMEDPRRAGKMLMTLLPLRQSTKAVQHFYNIKLEBK 443

Db 361 DVEAVOKLQDYLHEALQDYEAQGHMEDPRRAGKMLMTLLPLRQSTKAVQHFYNIKLEBK 420

QY 444 VPMHGLFLEMLEAKV 458

Db 421 VPMHGLFLEMLEAKV 435

RESULT 13

AAV83824

ID AAV83824 standard; protein; 418 AA.

XX AC AAV83824;

XX DE 05-JUN-2000 (first entry)

XX DE Human nuclear receptor nNR2-1 variant protein.

XX KM Human nuclear receptor protein 2-1; nNR2-1; physiological function; cell development and differentiation controller; gene expression.

XX OS Homo sapiens.

XX EN US6054295-A.

XX PD 25-APR-2000.

XX PF 26-AUG-1998; 98US-00141000.

XX PR 27-AUG-1997; 97US-0057090P.

XX PR 21-OCT-1997; 97US-0062922P.

XX PR 19-MAR-1998; 98US-0078633P.

XX PA (MERI) MERCK & CO INC.

XX PI Chen F;

XX DR WPI: 2000-328352/28.

XX DR N-PSDB; AAV09802.

XX PT New polynucleotide encoding human nuclear receptor protein-1 is useful for screening compounds that acts as modulators of cell differentiation, cell development and physiological function.

XX PS Disclosure; Fig 8; 58pp; English.

XX CC This sequence represents the variant human nuclear receptor protein 2-1 (nNR2-1). The coding gene differs from the nNR2 gene (AA09801) by the insertion of 2 nucleotides at position 1352. The insertion results in the shifting of the reading frame and the creation of a termination codon 33 nucleotides downstream of the insertion site. The translated protein has a C-terminal truncation. The nNR proteins are cell development and differentiation controllers which are useful for screening compounds that act as modulators of cell differentiation, cell development and physiological functions or for treating and diagnosing disorders associated with altered expression of the DNA

XX CC Sequence 418 AA;

QY Query Match 89.6%; Score 2139; DB 3; Length 418;

Db Best Local Similarity 100.0%; Pred. No. 3e-193;

Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVSELCPRESFSLHREELCRMKNKDHIDSSGCSFKTPRSPASITDSVNHHSPPG 60

Db 1 MSVSELCPRESFSLHREELCRMKNKDHIDSSGCSFKTPRSPASITDSVNHHSPPG 60

QY 61 SSDASGYSSTNMGNHONGDPSPLYPASAPILGGSGVVRKLYDDCSSTIYEDPOTKCEYML 120

Db 61 SSDASGYSSTNMGNHONGDPSPLYPASAPILGGSGVVRKLYDDCSSTIYEDPOTKCEYML 120

QY 121 NSMFKRLCTVCGDIASGHHGVASCEACKAFKKTIQGNITYSCPATNECITKRRKSC 180

Db 121 NSMFKRLCTVCGDIASGHHGVASCEACKAFKKTIQGNITYSOPATNECITKRRKSC 180

QY 181 QACRFMKCLKYGMKEGRLRVRGGRQYKRRIDAENSPLYNQLOVPAPKPYNKITVSH 240

Db 181 QACRFMKCLKYGMKEGRLRVRGGRQYKRRIDAENSPLYNQLOVPAPKPYNKITVSH 240

QY 241 LIVAPEKIVAMPDPTVPDSDIKALTTLLCDLADRELVIIMAKHIPPSTLSLADQMSL 300

Db 241 LIVAPEKIVAMPDPTVPDSDIKALTTLLCDLADRELVIIMAKHIPPSTLSLADQMSL 300

QY 301 LOSAMEILLIGVYRSLSPFEDELYADYIMDEDSQSLAGLIDLNNAIIQLVKKYKSMK 360

Db 301 LOSAMEILLIGVYRSLSPFEDELYADYIMDEDSQSLAGLIDLNNAIIQLVKKYKSMK 360

QY 361 LEKEEFVTLKAIALANDSMHIEDVEAVOKLQDYLHEALQDYEAQGHME 409

Db 361 LEKEEFVTLKAIALANDSMHIEDVEAVOKLQDYLHEALQDYEAQGHME 409

RESULT 14

AAV03838

ID AAV03838 standard; protein; 418 AA.

XX AC AAV03838;

XX DE 15-JUN-1999 (first entry)

XX DE Human nuclear receptor protein nNR2-1.

XX KM Nuclear receptor; nNR1; nNR2; cell differentiation; human.

XX OS Homo sapiens.

XX EN WO9910367-A1.

XX PD 04-MAR-1999.

XX PF 27-AUG-1998; 96WO-US017826.

XX PR 27-AUG-1997; 97US-0057090P.

XX PR 21-OCT-1997; 97US-0062902P.

XX PR 19-MAR-1998; 98US-0078633P.

XX PA (MERI) MERCK & CO INC.

XX PI Chen F;

XX DR WPI: 1999-190586/16.

XX DR N-PSDB; AAX32267.

XX PT Newly purified DNA polynucleotides encoding human nuclear trans-acting receptor proteins - useful in the diagnosis, treatment and prophylaxis of cell differentiation, development and physiological function.

XX PS Disclosure; Fig 8; 82pp; English.

XX CC The invention relates to DNA molecules encoding human nuclear receptor (nNR) proteins nNR1 and nNR2. The nNR proteins (including mutants and/or fragments) form pharmaceutical compositions that are useful in the diagnosis, treatment and prophylaxis of cell differentiation, development and physiological function. The proteins are also useful for identifying downstream target genes and ligands regulating their activity. In particular, fusion constructs (especially glutathione S-transferase (GST) nNR1 and/or GST-nNR2) expressing fusion proteins are useful in screening for (ant)agonists that are useful as modulators in cell differentiation, cell development and physiological function. Antibodies for the nNR proteins are useful for measuring the levels of these proteins. The present sequence represents a human nNR2-1 protein

XX CC Sequence 418 AA;

Query Match 88.9%; Score 2122; DB 2; Length 418;
 Best Local Similarity 99.5%; Pred. No. 1.2e-121;
 Matches 407; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSVELCLPESFSLHYEBELICRMSNKHIDSSCSFITEPSPASLTDSVNHSPGG 60
 DB 1 MDSVELCLPESFSLHYEBELICRMSNKHIDSSCSFITEPSPASLTDSVNHSPGG 60

QY 61 SSDSGSYSTSMNGHONGJLSPPLYPAPILIGSGGPVKLYDDCSSTIVEDPQKCEYML 120
 DB 61 SSDSGSYSTSMNGHONGJLSPPLYPAPILIGSGGPVKLYDDCSSTIVEDPQKCEYML 120

QY 121 NSMFKRLCLVCGDIASGYHYVASCEACAFKRTIQGNIEYSCPATNECEITRRRASC 180
 DB 121 NSMFKRLCLVCGDIASGYHYVASCEACAFKRTIQGNIEYSCPATNECEITRRRASC 180

QY 181 QACFPKCLKVGMLKEGVRILDRVGRGKRYKRIIDANSFYLPOLVOPAKKPYNKIVSH 240
 DB 181 QACFPKCLKVGMLKEGVRILDRVGRGKRYKRIIDANSFYLPOLVOPAKKPYNKIVSH 240

QY 241 LTVAEPEKIYAMPPTVPDSDIKALITLCLADREIVYIIGMAKHIGESTLSIADQMSL 300
 DB 241 LTVAEPEKIYAMPPTVPDSDIKALITLCLADREIVYIIGMAKHIGESTLSIADQMSL 300

QY 301 LOSAMEILILGVYRSLSFEDELVYADYIMEDOSKLAGLIDNNAILQLVKRYKSMK 360
 DB 301 LOSAMEILILGVYRSLSFEDELVYADYIMEDOSKLAGLIDNNAILQLVKRYKSMK 360

QY 361 LEKEEFTLTKAIALANDSMHIEVEAVOKLQVYLHEALQDYEAQOME 409
 DB 361 LEKEEFTLTKAIALANDSMHIEVEAVOKLQVYLHEALQDYEAQOME 409

RESULT 15
 ADD69753 standard; protein; 396 AA.
 AC ADD69753;
 DT 15-JAN-2004 (first entry)
 DE Human ERR gamma 3-related protein - SEQ ID 2.
 XX
 KW nuclear receptor; ERR gamma 3; oestrogen receptor-related receptor;
 KM oestrogen receptor; ER; thyroid hormone; TR; human.
 XX
 OS Homo sapiens.
 XX
 FN WC2003080831-A1.
 PD 02-OCT-2003.
 XX
 PR 25-MAR-2003; 2003WC-JP003611.
 XX
 PR 25-MAR-2002; 2002JP-00084560.
 XX
 PA (FUJI) FUJISAWA PHARM CO LTD.
 PI
 PI Koji H, Tajima K, Fukagawa M, Nishimura S, Isogai T;
 DR WPI, 2003-779262/73.
 DR N-PSDB; ADD69752.
 XX
 PT Polynucleotides encoding nuclear receptors, and the encoded proteins,
 PT useful as diagnostic agents, and for identification of agents that affect
 PT receptor activity.
 PS
 PS Claim 1; SEQ ID NO 2; 148bp; Japanese.
 CC
 CC The invention relates to novel nuclear receptor ERR (oestrogen receptor-
 CC related receptor) gamma 3 polynucleotides. The polynucleotides of the
 CC invention may be useful for diagnosis of disorders caused by abnormal

CC nuclear receptor activity, particularly those related to abnormal
 CC oestrogen receptor (ER), ERR or thyroid hormone receptor (TR) activity.
 CC Furthermore, the polynucleotides and proteins may be useful for
 CC evaluating agents that affect the activity of nuclear receptors. The
 CC current sequence is that of the human ERR gamma 3-related protein of the
 CC invention.
 XX
 SQ Sequence 396 AA;

Query Match 84.4%; Score 2016.5; DB 7; Length 396;
 Best Local Similarity 90.8%; Pred. No. 1.1e-161;
 Matches 395; Conservative 0; Mismatches 1; Indels 39; Gaps 1;

QY 24 MSNKHIDSSCSFITEPSPASLTDSVNHSPGSSDASGSYSTSMNGHONGJLSP 83
 DB 1 MSNKHIDSSCSFITEPSPASLTDSVNHSPGSSDASGSYSTSMNGHONGJLSP 83

QY 84 LYPAPILIGSGGPVKLYDDCSSTIVEDPQKCEYMLNSMFKRLCLVCGDIASGYHYGVA 143
 DB 61 LYPAPILIGSGGPVKLYDDCSSTIVEDPQKCEYMLNSMFKRLCLVCGDIASGYHYGVA 120

QY 144 SCEACAFKRTIQGNIEYSCPATNECEITRRRASCQACFPKCLKVGMLKEGVRILDRV 203
 DB 121 SCEACAFKRTIQ-----GVRILDRV 141

QY 204 RGRGQYKRYKRIIDANSFYLPOLVOPAKKPYNKIVSHLTVAEPEKIYAMPPTVPDSDIK 263
 DB 142 RGRGQYKRYKRIIDANSFYLPOLVOPAKKPYNKIVSHLTVAEPEKIYAMPPTVPDSDIK 201

QY 264 ALITLCLADREIVYIIGMAKHIGESTLSIADQMSLLOSAMEILILGVYRSLSFEDE 323
 DB 202 ALITLCLADREIVYIIGMAKHIGESTLSIADQMSLLOSAMEILILGVYRSLSFEDE 261

QY 324 LTVADYIMEDOSKLAGLIDNNAILQLVKRYKSMLEKEEFTLTKAIALANDSMHIE 383
 DB 262 LTVADYIMEDOSKLAGLIDNNAILQLVKRYKSMLEKEEFTLTKAIALANDSMHIE 321

QY 384 DVEAVOKLQVYLHEALQDYEAQOMEDPRRAGKMLMTLPILRQSTYKAVGHYNIKLEGK 443
 DB 322 DVEAVOKLQVYLHEALQDYEAQOMEDPRRAGKMLMTLPILRQSTYKAVGHYNIKLEGK 381

QY 444 VPMHKLFLFEMLEAKV 458
 DB 382 VPMHKLFLFEMLEAKV 396

Search completed: August 3, 2004, 12:40:31
 Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:38:36 ; Search time 17 Seconds
(without alignments)
2591.513 Million cell updates/sec

Title: US-10-054-841-4

Perfect score: 2388

Sequence: 1 MDSVELCLPESFSLHYEBEL.....KLEGVPMKFLFLEAKV 458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1784	74.7	433	2 B29345	steroid hormone re
2	1748	73.2	433	2 S58087	estrogen receptor
3	1223.5	51.2	589	1 A29345	steroid hormone re
4	718	30.1	589	1 ORCHE	estrogen receptor
5	716.5	30.0	600	1 ORRTE	estrogen receptor
6	713.5	29.9	595	2	estradiol receptor
7	710	29.7	599	1 ORMBE	estrogen receptor
8	702	29.4	595	1 ORHUE	estrogen receptor
9	694	29.1	586	1 ORLJE	estrogen receptor
10	674	28.2	535	2 S58224	estrogen receptor
11	674	28.2	620	2 T10423	estrogen receptor
12	666.5	27.9	574	2 A37197	estrogen receptor
13	662	27.7	701	2 S64737	80K estrogen recep
14	648	27.1	530	2 JCS939	estrogen receptor
15	644.5	27.0	477	2 S71400	estrogen receptor
16	634	26.5	503	2 JMO046	estrogen receptor
17	623.5	26.1	441	2 T50515	retinoid X recepto
18	609.5	25.5	470	2 D41977	retinoid receptor
19	600	25.1	488	2 C41977	retinoid receptor
20	598	25.0	467	2 A43781	retinoid X-recepto
21	596.5	25.0	467	2 S26668	retinoid acid rece
22	591.5	24.8	467	2 A47278	retinoid X recepto
23	589	24.7	462	1 S09592	retinoid X recepto
24	577.5	24.2	463	2 B41727	retinoid X recepto
25	573	24.1	379	2 T50514	retinoid X recepto
26	572.5	24.0	533	2 S37781	retinoid acid rece
27	572.5	24.0	461	2 S26670	retinoid acid rece
28	561	23.5	451	2 A41651	retinoid acid rece
29	560	23.5	446	2 A34418	H-2 region II bind

30	560	23.5	448	2 D41727	retinoid X recepto
31	560	23.5	452	2 S47633	RXR protein Afri
32	560	23.5	520	2 B47718	RXR-beta1 isoform
33	553.5	23.2	410	2 S26669	retinoid acid rece
34	551	23.1	422	2 T50516	retinoid X recepto
35	550.5	23.1	438	2 T50517	retinoid X recepto
36	545	22.8	476	2 B41977	retinoid acid rece
37	528.5	22.1	458	2 A34714	retinoid acid rece
38	527	22.1	454	1 A33903	retinoid acid rece
39	525	22.0	560	2 S27874	steroid hormone re
40	524	21.9	454	2 S06124	retinoid acid rece
41	522	21.9	442	2 A38592	retinoid acid rece
42	515.5	21.6	403	2 S35334	steroid receptor p
43	514	21.5	543	2 A32693	steroid receptor p
44	511	21.4	458	2 S06123	retinoid acid rece
45	509	21.3	443	1 C35991	retinoid acid rece

ALIGNMENTS

RESULT 1

B29345
steroid hormone receptor ERR2 precursor - human
C/Species: Homo sapiens (man)
C/Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 20-Sep-1999
C/Accession: B29345
R/Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.
Nature 331, 91-94, 1988
A/Title: Identification of a new class of steroid hormone receptors.
A/Reference number: A29345; PMID:88122546; PMID:3267207
A/Accession: B29345
A/Molecule type: mRNA
A/Residues: 1-433 <GIG>
A/Cross-references: GB:X51417; NID:G36610; PIDN:CA35779.1; PID:G36611; GB:Y00230
C/Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C/Keywords: DNA binding; steroid hormone receptor; transcription regulation; zinc finger
F:101-352/Domain: erbA transforming protein homology <ERRA>
F:103-123/Region: zinc finger
F:139-163/Region: zinc finger

Query Match	74.7%	Score 1784	DB 2	Length 433
Best Local Similarity	76.6%	Pred. No. 5.5e-124		
Matches 333	Conservative 51	Mismatches 49	Indels 2	Gaps 1
QY	24	MSNKRHIDSSGSSFIKTEPSSPASLTDSVNHHSFGSSSDASGSSTMNHGNDSP	83	
DB	1	MSEDRHIDSSGSSFIKTEPSSPSGIDALSHHSFGSSSDASGFGALGTNGLDSDP	60	
QY	84	LYPSAPILGSGGPKLYDDCSSTIYEDPOTKCYMLNSMPKLCVCGDILAGYHGYVA	143	
DB	61	MFAGAGL--GGNPCRSYEDCTSGIMEDSAIKCEYMLNIPKRLCVCGDILAGYHGYVA	118	
QY	144	SCBCKAFPRRTIQNGIIEYSCPATNCEITKRRRSQACREPKLVKGLKGVRLDRV	203	
DB	119	SCBCKAFPRRTIQNGIIEYSCPATNCEITKRRRSQACREPKLVKGLKGVRLDRV	178	
QY	204	RGRQRKRRIRDAENSPYINPOLVPAKKRYNKIVSHLVAEBEKIYAMDPPTVPSDIK	263	
DB	179	RGRQRKRRIRDAENSPYINPOLVPAKKRYNKIVSHLVAEBEKIYAMDPPTVPSDIK	238	
QY	264	ALTTTLDLADRELIVLIGAKKHIPGFSSTLSADOMSLQSANWEIILIGVYSLSPDE	323	
DB	239	ALTTTLDLADRELIVLIGAKKHIPGFSSTLSADOMSLQSANWEIILIGVYSLSPDDK	298	
QY	324	LYVADYIMDEDSKLAGLIDLNNAITQVKKYKSMLEKEEFTVTLKALANDSMHIE	383	
DB	299	LAVADYIMDEDSKLAGLIDLNNAITQVKKYKSMLEKEEFTVTLKALANDSMHIE	358	
QY	384	DVEAVQKLDVHEALQDYEAQGHMEDPRRAGMMLTPLLRTSTYKAVOHFNINIKLEGK	443	
DB	359	NIEAVQKLDVHEALQDYEAQGHMEDPRRAGMMLTPLLRTSTYKAVOHFNINIKLEGK	418	

QY 444 VPMKLTLEMLEAKV 458
 |||||
 Db 419 VPMKLTLEMLEAKV 433

RESULT 2

estrogen receptor related protein - mouse

C/Species: Mus musculus (house mouse)

C/Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999

C/Accession: S58087

R/Pettersson, K.; Svensson, K.; Mattsson, R.; Carlsson, B.; Ohlsson, R.; Berkenstam, A.

submitted to the EMBL Data Library, July 1995

A/Description: Expression of a novel member of estrogen response element-binding nuclear

A/Reference number: S58087

A/Accession: S58087

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-433 <PEP>

A/Cross-references: EMBL:X89594; NID:g91485; PDB:CAA61755.1; PID:g914856

C/Superfamily: unassigned exa-related proteins; exa transforming protein homology

C/Keywords: steroid hormone receptor; zinc finger

F:101-352/Domain: exa transforming protein homology <ERBA>

Query Match 73.2%; Score 1748; DB 2; Length 433;
 Best Local Similarity 74.9%; Pred. No. 2.5e-121;

Matches 326; Conservative 54; Mismatches 53; Indels 2; Gaps 1;

QY 24 MSNDRHIDSSCGSFITKEPSSPASLDSVNHHSPGSSSDASGSYSTMGCHQGLDSP 83
 |||||
 Db 1 MSSDRHIGSSCGSFITKEPSSPASLDSVNHHSPGSSSDASGSYSTMGCHQGLDSP 60
 |||||
 QY 84 LYPAPILGSGGPRKLYDDCSSTIVEDPQKCEYMSMPKLCIVCGDIASGVHYVA 143
 |||||
 Db 61 MPAGAGL--GNPORKSYEDCTSGIMEDSAIKCEYMNATPKRLCLVCGDIASGVHYVA 118
 |||||
 QY 144 SCEACKAFKRTTIOGNETEYSCPATNECEITRRRRKSCQACRFMKCLKVMKEGVRLDRV 203
 |||||
 Db 119 SCEACKAFKRTTIOGNETEYSCPATNECEITRRRRKSCQACRFMKCLKVMKEGVRLDRV 178
 |||||
 QY 204 RGGQKRRKRRIDANSPTLPOLVQPAKPEYNTKIVSHLVAPEKTYAMPPTVPDSDIK 263
 |||||
 Db 179 RGGQKRRKRRIDANSPTLPOLVQPAKPEYNTKIVSHLVAPEKTYAMPPTVPDSDIK 238
 |||||
 QY 264 ALTLICLDADDELVIITGMAGHIPGFTSLADQMSILIOSAMRLLITIGVYRSISPEDE 323
 |||||
 Db 239 ALTLICLDADDELVIITGMAGHIPGFTSLADQMSILIOSAMRLLITIGVYRSISPEDE 298
 |||||
 QY 324 LVVADYIMDEDSQKLAGLIDLNNAIQLVKKYSMKLEKEEFYTLKALALANSDSMHI 383
 |||||
 Db 299 LAVADYIMDEDSQKLAGLIDLNNAIQLVKKYSMKLEKEEFYTLKALALANSDSMHI 358
 |||||
 QY 384 DVEAVQKLDVHLALQDYEGQMEDPRRAGKMLMTPLRLQSTRAVQHFYNIKLBGK 443
 |||||
 Db 359 NLEAVQKLDVHLALQDYEGQMEDPRRAGKMLMTPLRLQSTRAVQHFYNIKLBGK 418
 |||||
 QY 444 VPMKLTLEMLEAKV 458
 |||||
 Db 419 VPMKLTLEMLEAKV 433

RESULT 3

steroid hormone receptor ERRT precursor - human

C/Species: Homo sapiens (man)

C/Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 20-Sep-1999

C/Accession: A29345; A49074

R/Giguere, V.; Yang, N.; Segui, P.; Evans, R.M.

Nature 331, 91-94, 1988

A/Title: Identification of a new class of steroid hormone receptors.

A/Reference number: A29345; MUID:86122546; PMID:3267207

A/Accession: A29345

A/Molecule type: mRNA
 A/Residues: 1-521 <GIG>
 A/Cross-references: EMBL:X51416; NID:g36608; PDB:CAA35778.1; PID:g36609; EMBL:Y00290
 R/Wiley, S.R.; Kraus, R.J.; Zuo, F.; Murray, E.E.; Lortitz, K.; Mertz, J.E.
 Genes Dev. 7, 2206-2219, 1993

A/Title: SV40 early-to-late switch involves titration of cellular transcriptional repres

A/Reference number: A49074; MUID:94040741; PMID:8224847

A/Accession: A49074

A/Status: preliminary

A/Molecule type: protein

A/Residues: 166-169, 'X', '171-173 <WIL>

C/Superfamily: unassigned exa-related proteins; exa transforming protein homology

C/Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; zi

F:174-434/Domain: exa transforming protein homology <ERBA>

F:176-196/Region: zinc finger

F:172-236/Region: zinc finger

Query Match 51.2%; Score 1223.5; DB 2; Length 521;
 Best Local Similarity 54.0%; Pred. No. 1.7e-82;
 Matches 252; Conservative 66; Mismatches 84; Indels 65; Gaps 9;

QY 22 CRMNKRHIDSSCGS-----FKTEPSSPASLDSVNHHSPGSSD-----A 64
 |||||
 Db 85 CPTALPEPQVTSAMSSQVVGIEPIYIKAPASP-----DSFKGSSSETETEPVALA 135
 |||||
 QY 65 SG-SYSTMGCHQGLDSPPIYPAPILGSGGPRKLYDDCSSTIVEDPQKCEYMSMPKLCIVCGDIASGVHYVA 123
 |||||
 Db 136 PGPAPTRCLPQHKEEDGE-----GAGP-----GEQGGKLVLSL 171
 |||||
 QY 124 PKRLCLVCGDIASGVHYVASCCEAKAFKRTTIOGNETEYSCPATNECEITRRRRKSCQAC 183
 |||||
 Db 172 PKRLCLVCGDIASGVHYVASCCEAKAFKRTTIOGNETEYSCPATNECEITRRRRKSCQAC 231
 |||||
 QY 184 RPKCLVGMKKEGVRLDRVGRQKYYRRIDANSPTLPOLVQ-----AKKY 234
 |||||
 Db 232 RFTKCLRVGMKKEGVRLDRVGRQKYYRRIDANSPTLPOLVQ-----AKKY 291
 |||||
 QY 235 NKIVSHLVAPEKTYAMPPTVPDSDIKALTTICLDADDELVIITGMAGHIPGFTSL 294
 |||||
 Db 292 NALVSHLVAPEKTYAMPPTVPDSDIKALTTICLDADDELVIITGMAGHIPGFTSL 351
 |||||
 QY 295 ADQMSILIOSAMRLLITIGVYRSISPEDELYADYIMDEDSQKLAGLIDLNNAIQLVYK 354
 |||||
 Db 352 SDQMSIVQSWMEVLVGVAGRSPLDDELAFADVDVDESGARAAGELGALLQLVYR 411
 |||||
 QY 355 KYKSMKLEKEEFYTLKALALANSDSMHI EDYEA V-QKLQVYHLALQDYEGQ----HME 409
 |||||
 Db 412 RLQALRLEREYVYTLKALALANSDSMHI EDEPR LMSCEKILHALLLEYEAGRAPGGGA 471
 |||||
 QY 410 DPRRAGKMLMTPLRLQSTRAVQHFYNIKLBGKVPKLTLEMLEA 456
 |||||
 Db 472 ERRRAGKMLMTPLRLQSTRAVQHFYNIKLBGKVPKLTLEMLEA 518
 |||||

RESULT 4

ORCHE

estrogen receptor - chicken

C/Species: Gallus gallus (chicken)

C/Date: 06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change 22-Jun-1999

C/Accession: A40914; S07192

R/McConnell, B.L.; McDowell, D.P.; Conneely, O.M.; Schultz, T.Z.; Greene, G.L.; O'Malley, J.

Mol. Endocrinol. 1, 25-35, 1987

A/Title: Structural organization and regulation of the chicken estrogen receptor.

A/Reference number: A40914; MUID:88318621; PMID:2901032

A/Accession: A40914

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-589 <MAX>

R/Kunst, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.; Bornert, J.M.; Chambon, P.

EMBO J. 5, 891-897, 1986

A/Title: The chicken estrogen receptor sequence: homology with v-erbA and the human oes

A/Reference number: S07192; MUID:86247578; PMID:3755102

A/Accession: S07192

A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-256, 'E', 258-589 <KRU>
A:Cross-references: EMBL:X03805; NID:g63378; PIDN:CAA27433.1; PID:g63380
C:Comment: The steroid hormones and their receptors are involved in the regulation of eu
C:Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly
complex appears to recognize discrete DNA sequences upstream of transcriptional start sit
C:Superfamily: estrogen receptor; e2bA transforming protein homology
C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone recep
F:1-13/Domain: amino-terminal <NH2>
F:174-265/Domain: DNA binding #status predicted <DNA>
F:177-450/Domain: e2bA transforming protein homology <ERBA>
F:117-200/Region: zinc finger CCCC motif
F:213-235/Region: zinc finger CCCC motif
F:250-265/Region: nuclear location signal
F:294-546/Domain: steroid binding #status predicted <STB>
F:179, 182, 196, 199/Binding site: zinc (Cys) #status predicted
F:215, 221, 231, 234/Binding site: zinc (Cys) #status predicted
F:230, 299/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 30.1%; Score 718; DB 1; Length 589;
Best Local Similarity 36.8%; Pred. No. 4,3e-45;
Matches 170; Conservative 90; Mismatches 134; Indels 68; Gaps 14;

QY 38 FIKTESSPASLSDSVNHS---PGSSDASGSYSTNMHONGLDSPLY--PSAPILG 92
DB 104 FQGTAR---QUSPFIHHSQVYLYLENGSF---GMRA--APPAFRPSS--- 148
QY 93 GSGPVEKLYDDSSITVEDPQTKCE---YMLSMERKLCIVCGDIASGVHYGASCEAC 148
DB 149 -----DNRRHSIREMSSSTNEKGSLSMESTKETRCAVONDVASGHVYVMSCEGC 199
QY 149 KAFPEKTIQNTIEYSCPATNECEITRRRRKSCQACFPKCLKYGMLEKGVRLDRVGR- 207
DB 200 KAFPEKSIQHDYMCAPATNQCTIDKRRKSCQACSLRKYCYEVMKGGIRKOR--RGEM 258
QY 208 QYKRRIDANSP-----YINPOLVQAKK-----PYNKIVSHLIVAB 246
DB 259 MOKRREBQSDSNGASSTELRAPLTMSPVAVKNNKNSPALSLTAEQMSALLAEAP 318
QY 247 EKLYAPDPPTVPDSIKALITLDCDLADRELVIIGAKHIPGSTSLADQMSILQSAWM 306
DB 319 PLYSVYDPNRPENASMTLTNLADRELVIHAKVPGVVDLTLDQVHLECAHL 378
QY 307 EIIILGVYRSLEFEDIVAYDDYIMDEQSK--LAGLIDLNNAIIQVKKYKSMKLEKE 365
DB 379 EIIIMGLVMSMEHPKLTLPAPLTLDRQKCEVGEVETFDMLLTAAFRFMNMQEE 438
QY 366 FYTLKALALAN-----SDSMHIDVEAVQKLDQVHEALQDYAGQHM--EDPRRA 414
DB 439 FVCLKSIILINSGVYTLSTLTKSLBERDYIHRVLDKIDTTLIHLMAKSGSLDQOHRRL 498
QY 415 GKMLMTLPRLRQSTKAVOHFYNIKLEGVPMHKLFLMLEA 456
DB 499 AOLILLSHIRHMSNKGMEHLVNMCKKNVPLXDLLEMLDA 540

RESULT 5
ORRTE
estrogen receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Sep-1991 #sequence_revision 14-Jul-1994 #text_change 22-Jun-1999
C:Accession: S07379; S16731
C:Koike, S.; Sakai, M.; Muramatsu, M.
Nucleic Acids Res. 15, 2499-2513, 1987
A:Title: Molecular cloning and characterization of rat estrogen receptor cDNA.
A:Reference number: S07379; PMID:87174780; PMID:3031601
A:Accession: S07379
A:Molecule type: mRNA
A:Residues: 1-600 <ROI>
A:Cross-references: EMBL:Y00102; NID:g56110; PIDN:CAA68287.1; PID:g56111
R:Maggi, A.M.A.
submitted to the EMBL Data Library, June 1991

A:Reference number: S16731
A:Accession: S16731
A:Molecule type: mRNA
A:Residues: 1-487, 'T', 489-600 <MAG>
A:Cross-references: EMBL:X61098; NID:g56120; PIDN:CAA43411.1; PID:g56121
C:Comment: The steroid hormones and their receptors are involved in the regulation of eu
C:Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly
complex appears to recognize discrete DNA sequences upstream of transcriptional start sit
C:Superfamily: estrogen receptor; e2bA transforming protein homology
C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone recep
F:1-184/Domain: amino-terminal <NH2>
F:185-276/Domain: DNA binding #status predicted <DNA>
F:188-461/Domain: e2bA transforming protein homology <ERBA>
F:188-211/Region: zinc finger CCCC motif
F:224-246/Region: zinc finger CCCC motif
F:261-276/Region: nuclear location signal
F:305-557/Domain: steroid binding #status predicted <STB>
F:190, 193, 207, 210/Binding site: zinc (Cys) #status predicted
F:226, 232, 242, 245/Binding site: zinc (Cys) #status predicted
F:241, 310/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 30.0%; Score 716.5; DB 1; Length 600;
Best Local Similarity 39.2%; Pred. No. 5,7e-45;
Matches 168; Conservative 82; Mismatches 128; Indels 51; Gaps 13;

QY 73 NGHONGLDSPLY---PSAPILGSGP---VRKLYDDCSITVEDPQTKCE---YMLN 121
DB 129 HGHO-----VPLYENEPVAVAVDTGPPAPFRKSNDSNRRONGRERLSSSEKNNIMES 183
QY 122 SMPKRLCIYCGDIASGVHYGASCEACKAPFRTIQNTIEYSCPATNECEITRRRRKSCQ 181
DB 184 AKRTYCAVNDYASGYHYGWSCEGCKAPFRKSIQHDYMCAPATNQCTIDKRRKSCQ 243
QY 182 ACFPMKCLYGMLEKGVRLDRVGRQ--KYKR-RIDAEN-----SPYINPOL 226
DB 244 ACRLRKYCVGMKGGIRKOR--RGGMILKHKKQDDLEGRNEMGSGDRAANLWSPV 302
QY 227 VQPAK-----PYNKIVSHLIVABPEKIYAMPPTVPDSIKALITLDCDLADRELVI 279
DB 303 IKHYTKNSPALSLTDQMSVALLDAEPPLIYSEYDPSRPFSEASMMGLTLNLADRELIVM 362
QY 280 IGMAKHIPGSTSLADQMSILQSAAMELIIIGVYRSLEFEDIVAYDDYIMDEQSK- 338
DB 363 IYMAKVPGPQDNLNDQVHLECAWLITLMGLVMSMEHPKLTLPAPLTLDRQKGC 422
QY 339 LAGLIDLNNAIIQVKKYKSMKLEKEFYTLKALALAN-----SDSMHIDVEAVQK 390
DB 423 VEGWEIIPMLLATSRPFEMNLQGEFEVCLKSIILINSGVYTLSTLTKSLKXKHHR 482
QY 391 LQDVLEHALQDY--EAGQMEBD--RRAGKMLTLPRLRQSTKAVOHFYNIKLEGVPMH 447
DB 483 VLDKINDTLIHLMAKAGLTLOOQHRRLAOLILLSHIRHMSNKGMEHLVNMCKKNVPLX 542
QY 448 KLFLEMLEA 456
DB 543 DLEMLDA 551

RESULT 6
147140
estradiol receptor - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Aug-1999
C:Accession: I47140; S66250; S32402
C:Bokenkamp, D.; Jungblut, P.W.; Thole, H.H.
Mol. Cell. Endocrinol. 104, 163-172, 1994
A:Title: The C-terminal half of the porcine estradiol receptor contains no post-translat
A:Reference number: I47140; PMID:95080454; PMID:7988744
A:Accession: I47140
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-595 <BOK>
A:Cross-references: EMBL:Z37167; NID:g587554; PIDN:CAA85524.1; PID:g587555

A:Experimental source: uterus
 R:Thole, H.H.; Maschler, I.; Jungblut, P.W.
 Eur. J. Biochem. 231, 510-516, 1995
 A:Title: Surface mapping of the ligand-filled C-terminal half of the porcine estradiol receptor
 A:Reference number: S66250; MUID:95361877; PMID:7635163
 A:Accession: S66250
 A:Molecule type: protein
 A:Residues: 297-307;310-313;320-323;329-332;337-340;417-420;466-473 <THO>
 A:Experimental source: uterus
 R:Thole, H.H.
 FEBS Lett. 320, 92-96, 1993
 A:Title: Assignment of the ligand binding site of the porcine estradiol receptor to the
 A:Reference number: S32402; MUID:93209384; PMID:8458437
 A:Accession: S32402
 A:Molecule type: protein
 A:Residues: 305-323 <THO>
 C:Superfamily: estrogen receptor; erba transforming protein homology
 C:Keywords: DNA binding; nucleus; phosphoprotein; steroid hormone receptor; transcriptic
 F:1-120/Domain: amino-terminal <NH2>
 F:121-239/Domain: DNA binding #status predicted <DNA>
 F:183-456/Domain: erba transforming protein homology <ERBA>
 F:185-205/Region: zinc finger CCCC motif
 F:221-245/Region: zinc finger CCCC motif
 F:256-271/Region: nuclear location signal
 F:300-595/Domain: steroid binding #status predicted <STB>
 F:185,188,202,205/Binding site: zinc (Cys) #status predicted
 F:221,227,237,240/Binding site: zinc (Cys) #status predicted
 F:236,305/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 29.9%; Score 713.5; DB 2; Length 595;
 Best Local Similarity 42.3%; Pred. No. 9.3e-45;
 Matches 159; Conservative 69; Mismatches 105; Indels 43; Gaps 10;

QY	119	MLNSMPKRLCLVCGDIALSGHYGYVASCACAFKRTIQGNIYSCPATNCEITKRRK	178
DB	176	MESAKETRYCAVNCNDYAGHYGYWSCGCAFFKRSIQGNDYVCPATNCTIDKRRK	235
QY	179	SCQACRFKCLKVGMLKEGVLDLRYRGGRQ-KYKR-RIDAEN-----SPYLN	223
DB	236	SCQACRLKCYEVGMKGGIRKDR-RGGRMLKHKQRDLGRNMGASGMRANLWPS	294
QY	224	POLVOPAKK-----PYNKIVSHLVAPEKTYAMPPTVPDSDIKALTYLCLDADEL	276
DB	295	PLIKHTKNSPVALSLTDQMSIALBAEPRIYSEYPTPLSPASMMGLITNLADREL	354
QY	277	VVIIGMAHIGFSTLSLADQMSLIQSAMWEILLGVYRSISFDELVYADYIMDEDO	336
DB	355	VHMINMARVPGFGLDLSHDVHLLECAMLEILMGLVRSMEHPGKLLFAPNILLDRNQ	414
QY	337	SK-LAGLIDLNNALIQLVKKYKSKMLEKEEFVTLKAIKALANS-----DSMH	381
DB	415	GKCEGWEIYEDMLATSSRFRMNLQGEFEVCLSKITILNSGVYTPPLSLSTLKSEEDH	474
QY	382	IEDVEAVOKLDVLEALQDYEAQOMEDP-RRAGKMLMTPLRLQSTKAVOHFYNTKL	440
DB	475	IHRV-LDKITDTLHMA--KAGLTLOOQHRRLAQLLILSHFRHMSKNGMEHLNMYKC	530
QY	441	EKGVPNKLPLEMDEA 456	
DB	531	KNVVPVLYDLLEMLDA 546	

RESULT 7
 QRMSE
 estrogen receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 06-May-1992 #sequence_revision 14-Jul-1994 #text_change 22-Jun-1999
 C:Accession: A40061
 R:White, R.; Lees, J.A.; Needham, M.; Ham, J.; Parker, M.
 Mol. Endocrinol. 1, 735-744, 1987
 A:Title: Structural organization and expression of the mouse estrogen receptor.
 A:Reference number: A40061; MUID:91042558; PMID:2484714
 A:Accession: A40061

A:Molecule type: mRNA
 A:Residues: 1-599 <WH>
 A:Cross-references: GB:M8651; NID:9193179; PIDN:AAA37580.1; PID:9193180
 C:Comment: The steroid hormones and their receptors are involved in the regulation of eu
 C:Comment: In the absence of ligand, steroid hormone receptors are thought to be weakly
 C:Comment: complex appears to recognize discrete DNA sequences upstream of transcriptional start sit
 C:Superfamily: estrogen receptor; erba transforming protein homology
 C:Keywords: DNA binding; nucleus; phosphoprotein; steroid binding; steroid hormone recepti
 F:1-183/Domain: amino-terminal <NH2>
 F:184-275/Domain: DNA binding #status predicted <DNA>
 F:187-460/Domain: erba transforming protein homology <ERBA>
 F:187-210/Region: zinc finger CCCC motif
 F:223-245/Region: zinc finger CCCC motif
 F:260-275/Region: nuclear location signal
 F:304-556/Domain: steroid binding #status predicted <STB>
 F:189,192,206,209/Binding site: zinc (Cys) #status predicted
 F:225,231,241,244/Binding site: zinc (Cys) #status predicted
 F:240,309/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 29.7%; Score 710; DB 1; Length 599;
 Best Local Similarity 41.8%; Pred. No. 1.7e-44;
 Matches 157; Conservative 72; Mismatches 104; Indels 43; Gaps 10;

QY	119	MLNSMPKRLCLVCGDIALSGHYGYVASCACAFKRTIQGNIYSCPATNCEITKRRK	178
DB	180	MESAKETRYCAVNCNDYAGHYGYWSCGCAFFKRSIQGNDYVCPATNCTIDKRRK	239
QY	179	SCQACRFKCLKVGMLKEGVLDLRYRGGRQ-KYKR-RIDAEN-----SPYLN	223
DB	240	SCQACRLKCYEVGMKGGIRKDR-RGGRMLKHKQRDLGRNMGASGMRANLWPS	298
QY	224	POLVOPAKK-----PYNKIVSHLVAPEKTYAMPPTVPDSDIKALTYLCLDADEL	276
DB	299	PLIKHTKNSPVALSLTDQMSIALBAEPRIYSEYPTPLSPASMMGLITNLADREL	358
QY	277	VVIIGMAHIGFSTLSLADQMSLIQSAMWEILLGVYRSISFDELVYADYIMDEDO	336
DB	359	VHMINMARVPGFGLDLSHDVHLLECAMLEILMGLVRSMEHPGKLLFAPNILLDRNQ	418
QY	337	SK-LAGLIDLNNALIQLVKKYKSKMLEKEEFVTLKAIKALANS-----DSMH	381
DB	419	GKCEGWEIYEDMLATSSRFRMNLQGEFEVCLSKITILNSGVYTPPLSLSTLKSEEDH	478
QY	382	IEDVEAVOKLDVLEALQDYEAQOMEDP-RRAGKMLMTPLRLQSTKAVOHFYNTKL	440
DB	479	IHRV-LDKITDTLHMA--KAGLTLOOQHRRLAQLLILSHFRHMSKNGMEHLNMYKC	534
QY	441	EKGVPNKLPLEMDEA 456	
DB	535	KNVVPVLYDLLEMLDA 550	

RESULT 8
 QRMSE
 estrogen receptor 1 - human
 N:Alternate names: ERI; estrogen receptor alpha
 C:Species: Homo sapiens (man)
 C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 22-Jun-1999
 C:Accession: A94284; A93376; A40021; S27123; S34000; A41925; B41925; A03244; C41925; D41
 R:Greene, G.L.; Gilna, P.; Waterfield, M.; Baker, A.; Hort, Y.; Shine, J.
 Science 231, 1150-1154, 1986
 A:Title: Sequence and expression of human estrogen receptor complementary DNA.
 A:Reference number: A94284; MUID:86122927; PMID:3753802
 A:Accession: A94284
 A:Molecule type: mRNA
 A:Residues: 1-595 <GR>
 A:Cross-references: GB:M12674; NID:9182192; PIDN:AAA52399.1; PID:9182193
 R:Green, S.; Walter, P.; Kumar, V.; Krust, A.; Bornert, J.M.; Argos, P.; Chambon, P.
 Nature 320, 134-139, 1986
 A:Title: Human estrogen receptor cDNA: sequence, expression and homology to v-erb-A.
 A:Reference number: A93376; MUID:86146892; PMID:3754034
 A:Accession: A93376
 A:Molecule type: mRNA

A:Reference number: A37197; MUID:91006824; PMID:2210031
 A:Accession: A37197
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-574 <PAK>
 A:Cross-references: GB:M31559
 R:Reikel, F.; Le Guellec, C.; Vaillant, C.; Le Roux, M.G.; Valotaire, Y.
 Mol. Endocrinol. 3, 44-51, 1989
 A:Title: Identification and estrogen induction of two estrogen receptors (ER) messenger
 A:Reference number: A40070; MUID:89127284; PMID:2315648
 A:Accession: A40070
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 150-574 <PA2>
 A:Cross-references: GB:M31559; NID:9213783; PIND:AAA9552.1; PID:9213784
 C:Superfamily: estrogen receptor; extra transforming protein homology
 C:Keywords: DNA binding, nucleus; steroid hormone receptor; transcription regulation; z
 F:145-421/Domain: erba transforming protein homology <ERBA>
 F:147-167/Region: zinc finger
 F:183-207/Region: zinc finger

Query Match 27.9%; Score 666.5; DB 2; Length 574;
 Best Local Similarity 35.0%; Pred. No. 2.6e-41;

Matches 167; Conservative 79; Mismatches 142; Indels 89; Gaps 17;

```

QY 38 FIKTEPSPASLIDSVNHSFGSSSDASGSY-----SSTMNGHNGIDSPILYPSAIIIG 93
DB 64 FVSSSHQSPQSPFL--HPSHHGLPSQSYLHNSST-----PLYRSSVYVINO 110
QY 94 -SGPRKLYDDCSSTIVEDPQ-----TKCEVNLNMPKRLCLVCGDIASGYHYGA 143
DB 111 LSAAEEKL-----CIASDROQSYSAAGSVNFEMANE--TRYCAVCSDFASGHYGV 162
QY 144 SCEACAPFKRTIOGNIIEYSCPATNECETTKRRKSCQACRFMKCLKVMKEGYELDRV 203
DB 163 SEEGKAPFKRSIQHNDYCPATNQCTDRNRKSCQACRLKCYEVAVGGRKRD-- 220
QY 204 RGR--OKYKRI-----DAE-----NSPYLN-----POLYQPAK 232
DB 221 RGRVLRMDKRYCGAPGRKPTVWSTGQRPQDGRNSSLNNGGGRGRTITPPEQ 280
QY 223 PYNKIVSLVAEPEKIYAMPDPTVPDSIDKALITLCLADRELVIIGAKHIFGSTL 292
DB 281 -----VLPLQGTPLCSRKQKAPRYTEVTMTTLTSMADKEIVHMLAKKVGPFEL 335
QY 293 SLADQMSLQSAWMEIILGVVYRSLFEDELVAADYIMEDQSK-LAGLIDNNALIQ 351
DB 336 SLHDVQLLESSWLEVMITGLWRSIHCPGKLIQAQDILIDRSBGDCVEGMAEIFDMLA 395
QY 352 LVKKYKMKLEKEFEVTLKAIALANS-----DSMHTEDVAVQKLDVLEAL-- 399
DB 396 TVSRGMLKLPREPVCAKAIILNPGAFSCNSVESLH--NSAVESMDNTIDALIH 453
QY 400 -QDYAGQHMEDPRRAGKMLTLLRLQSTKAVOHFNKILKGVPMKFLFEMLE 455
DB 454 HHSAGSVQOQPRROQLLTLISHIRMSNKGMEHLYSICKKKVPIYDILLEMID 510

```

RESULT 13

80K estrogen receptor - human
 C:Species: Homo sapiens (man)
 C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Aug-1999
 C:Accession: S64737
 R:Pink, J.T.; Wu, S.Q.; Wolf, D.M.; Billimoria, M.M.; Jordan, V.C.
 Nucleic Acids Res. 24, 962-969, 1996
 A:Title: A novel 80 kDa human estrogen receptor containing a duplication of exons 6 and
 A:Reference number: S64737; MUID:96174665; PMID:8600466
 A:Accession: S64737
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-701 <PIN>
 A:Cross-references: EMBL:U47678; NID:91197854; PIND:AAH00115.1; PID:91197855

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996
 C:Superfamily: estrogen receptor; extra transforming protein homology
 C:Keywords: steroid hormone receptor; zinc finger
 F:183-456/Domain: erba transforming protein homology <ERBA>
 F:148-562/Domain: erba transforming protein homology #status atypical <ERB2>

Query Match 27.7%; Score 662; DB 2; Length 701;
 Best Local Similarity 30.7%; Pred. No. 7.3e-41;

Matches 179; Conservative 80; Mismatches 132; Indels 192; Gaps 17;

```

QY 56 HSPGSSSDASGSSTSTMNGHNGIDS-PPL--YPS-----APILGSGSPYKRL 100
DB 80 YEGSEAAAFGS-----NGLGFPPLNSVSPSLMLLHPPLQSPFLQHGQGV 130
QY 101 Y--DDCSSTIVEDPQTKCEVNLN-----MKRLCLVCGD 133
DB 131 YLENESGTVYAEAGPRATYRRNSDRQGRRLASINDKGSMAEKETGYCAVND 190
QY 134 IASGYHYVASCEACAPFKRTIOGNIIEYSCPATNECETTKRRKSCQACRFMKCLKVM 193
DB 191 VASGYHYGVASCEGKAPFKRSIQHNDYCPATNQCTIDKRRKSCQACRLKCYEVGA 250
QY 194 LKEGRLDPRGRQ--KYYR-RIDAE-----SPYLPOLYQPAK-- 232
DB 251 MKGGRKOR--RGRMLKRRQRDDEGRGEVGSAGDMRAANIMPSPMLTKSKNSIALS 309
QY 233 -PYNKIVSLVAEPEKIYAMPDPTVPDSIDKALITLCLADRELVIIGAKHIFGST 291
DB 310 LTADQMSALDAEPILYSEYDPTFRPSRSMGLITNLADRELVHMLAKKRVGFVD 369
QY 232 ISLADQMSLQSAWMEIILGVVYRSLFEDELVAADYIMEDQSK-LAGLIDNNALIQ 350
DB 370 LTHQVHLLECAMLEIIMIGVWRSMEHPVTLFAPNILLDRNGKCYEVAVELFDM 429
QY 351 QLVKKYKMKLEKEFEVTLKAIALANS-----DSMH-----IED-- 384
DB 430 ATSSRFMMNLQGESEFVCLKSIILNSGYVTFSLSLKLEKDHIRVLDKITDILH 489
QY 385 -----VEAVQKLDY-- 394
DB 490 MAXKGLITQOQHRLAQLLILSHIRHRNQKCYEVAVELFDMILATSSRFMMNLQGE 549
QY 395 -----LHBAIDYEAQHN-----EDPR 413
DB 550 EFVCLKSIILNSGYVTFSLSLKLEKDHIRVLDKITDILHMAKAGITLQOQH 609
QY 414 AGKMLTLPRLRQSTKAVOHFNKILKGVPMKFLFEMLEA 456
DB 610 LAQLLILSHIRHMSNKGMEHLYSICKKNVPIYDILLEMIDA 652

```

RESULT 14

JC5939
 estrogen receptor beta - human
 C:Species: Homo sapiens (man)
 C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jul-2000
 C:Accession: JC5939
 R:Ogawa, S.; Inoue, S.; Watanabe, T.; Hiroi, H.; Orito, A.; Hosoi, T.; Ouchi, Y.; Muramat
 Biochem. Biophys. Res. Commun. 243, 122-126, 1998
 A:Title: The complete primary structure of human estrogen receptor beta (hERbeta) and its
 A:Reference number: JC5939; MUID:98139878; PMID:9473491
 A:Accession: JC5939
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-530 <OGA>
 A:Cross-references: DDBJ:AB006590; NID:92911151; PIND:BA24953.1; PID:92911152
 C:Superfamily: estrogen receptor; extra transforming protein homology
 F:147-408/Domain: erba transforming protein homology <ERB>

Query Match 27.1%; Score 648; DB 2; Length 530;
 Best Local Similarity 35.1%; Pred. No. 5.5e-40;

Matches 174; Conservative 92; Mismatches 145; Indels 84; Gaps 23;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:37:35 ; Search time 13 Seconds

(without alignments)
1834.472 Million cell updates/sec

Title: US-10-054-841-4

Perfect score: 2388

Sequence: 1 MDVSELCLPESFSIHVEEL.....KLEGVPMKIFLEMLKAV 458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	2388	100.0	458	ERR3_HUMAN	O75454 homo sapien
2	1799	75.3	500	ERR2_HUMAN	O95718 homo sapien
3	1784	74.7	433	ERR2_RAT	P11475 rattus norv
4	1748	73.2	433	ERR2_MOUSE	O61539 mus musculu
5	1258.5	52.7	519	ERR1_HUMAN	P11474 homo sapien
6	1253.5	52.5	462	ERR1_MOUSE	O08580 mus musculu
7	716.5	30.0	600	ERR1_RAT	P06212 rattus norv
8	715	29.9	589	ERR1_CHICK	O29040 sus scrofa
9	713.5	29.9	595	ERR1_PIG	O91250 poephilia gu
10	712.5	29.8	587	ERR1_PIG	O94235 mesocricetu
11	711.5	29.8	595	ERR1_PIG	P19785 mus musculu
12	710	29.7	599	ERR1_MOUSE	P03372 homo sapien
13	704	29.5	594	ERR1_HORSE	O94998 equus cabal
14	704	29.5	595	ERR1_HUMAN	P81559 xenopus lae
15	694	29.1	586	ERR1_XENLA	P57717 brachydanto
16	690	28.9	569	ERR1_BRARE	O9pvn9 sparus aura
17	687.5	28.8	579	ERR1_SPRAU	O9yha3 oreochromis
18	684.5	28.7	585	ERR1_ORENI	P16058 oncorhynch
19	682.5	28.6	622	ERR1_ONCMY	O42132 pacrus majo
20	678	28.4	581	ERR1_PAGMA	O9pve2 sturnus vul
21	676	28.3	554	ERR2_STYPU	O9yha2 ictalurus p
22	674.5	28.2	617	ERR1_ICTPU	P50242 salmo salar
23	674	28.2	535	ERR1_SALSA	O9p247 coryzias lat
24	674	28.2	620	ERR1_ORYLA	O9p241 coryzias lat
25	671.5	28.1	472	ERR1_ORYLA	O9p241 coryzias lat
26	666.5	27.9	583	ERR2_CHICK	O9p241 coryzias lat
27	661.5	27.7	583	ERR2_COTJA	O9p241 coryzias lat
28	661.5	27.7	583	ERR2_COTJA	O9p241 coryzias lat
29	655.5	27.4	530	ERR2_MITCN	P57781 microgonis
30	653.5	27.4	530	ERR2_RAT	O9p241 coryzias lat
31	653	27.3	568	ERR1_MITCN	P57781 microgonis
32	652.5	27.3	568	ERR2_ONCMY	P57782 oncorhynch
33	650	27.2	530	ERR2_SHEEP	O9p241 coryzias lat
				ERR2_MOUSE	O08537 mus musculu

RESULT 1	ERR3_HUMAN	STANDARD	PRT	458 AA.	
ID	ERR3_HUMAN	O75454; O96021; O9R1F3;			O92731 homo sapien
AC	O75454; O96021; O9R1F3;	15-JUL-1999 (Rel. 38, Created)			O9w659 carassius a
DT	15-JUL-1999 (Rel. 38, Created)	16-OCT-2001 (Rel. 40, Last sequence update)			O13012 anguilla ja
DT	16-OCT-2001 (Rel. 40, Last sequence update)	10-OCT-2003 (Rel. 42, Last annotation update)			O9xsb5 bos taurus
DE	Estrogen-related receptor gamma (Estrogen receptor related protein 3)	(ERR gamma-2)			O9yha3 microgonis
DE	(ERR gamma-2)	ESRRG OR NR3B3 OR ERGR2 OR ERK3 OR KIAA0832.			O9xsw2 sus scrofa
GN	ESRRG OR NR3B3 OR ERGR2 OR ERK3 OR KIAA0832.	Mus musculus (Mouse), and			O9yha2 ictalurus p
OS	Mus musculus (Mouse), and	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			O91ak1 carassius a
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	NCBI Taxid=9606, 10090;			O90416 brachydanto
OX	NCBI Taxid=9606, 10090;	[1]			
RN	[1]	SEQUENCE FROM N.A. (ISOFORM LONG).			
RP	SEQUENCE FROM N.A. (ISOFORM LONG).	SPECIES=Human; TISSUE=Brain;			
RX	MEDLINE=99173874; PubMed=10072763;	Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,			
RA	Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,	Lin Q., Caskey C.T.;			
RT	Identification of two hERR2-related novel nuclear receptors utilizing	bioinformatics and inverse PCR.";			
RL	Gene 228:101-109(1999).	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM LONG).	SPECIES=Human; TISSUE=Brain;			
RX	MEDLINE=99156230; PubMed=10048485;	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,			
RA	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;			
RT	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;	"Prediction of the coding sequences of unidentified human genes. XII.			
RL	"Prediction of the coding sequences of unidentified human genes. XII.	The complete sequences of 100 new cDNA clones from brain which code			
RP	The complete sequences of 100 new cDNA clones from brain which code	for large proteins in vitro.";			
RX	DNA Res. 5:355-364(1998).	[3]			
RA	SEQUENCE FROM N.A. (ISOFORM SHORT).	SPECIES=Human;			
RT	SPECIES=Human;	MEDLINE=99341123; PubMed=9676434;			
RL	MEDLINE=99341123; PubMed=9676434;	Eudy J.D., Yao S.F., Weston M.D., Ma-Edmonds M., Talmadge C.B.,			
RP	Eudy J.D., Yao S.F., Weston M.D., Ma-Edmonds M., Talmadge C.B.,	Cheng J.J., Kimberling W.J., Sumegi J.;			
RX	Cheng J.J., Kimberling W.J., Sumegi J.;	"Isolation of a gene encoding a novel member of the nuclear receptor			
RA	"Isolation of a gene encoding a novel member of the nuclear receptor	superfamily from the critical region of Usher syndrome type IIA at			
RT	superfamily from the critical region of Usher syndrome type IIA at	1q41.";			
RL	1q41."	Genomics 50:382-384(1998).			
RP	Genomics 50:382-384(1998).	[4]			
RX	SEQUENCE FROM N.A.	SPECIES=Mouse;			
RA	SPECIES=Mouse;	MEDLINE=99357798; PubMed=10428842;			
RT	MEDLINE=99357798; PubMed=10428842;	Hong H., Yang L., Stallcup M.R.;			
RL	Hong H., Yang L., Stallcup M.R.;	"Hormone-independent transcriptional activation and coactivator			
RP	"Hormone-independent transcriptional activation and coactivator	binding by novel orphan nuclear receptor ERR3.";			
RX	binding by novel orphan nuclear receptor ERR3.";	J. Biol. Chem. 274:22618-22626(1999).			
RA	J. Biol. Chem. 274:22618-22626(1999).	"- FUNCTION: ORPHAN RECEPTOR. BINDS SPECIFICALLY TO AN ESTROGEN			
RT	"- FUNCTION: ORPHAN RECEPTOR. BINDS SPECIFICALLY TO AN ESTROGEN	RESPONSE ELEMENT AND ACTIVATES REPORTER GENES CONTROLLED BY			
RL	RESPONSE ELEMENT AND ACTIVATES REPORTER GENES CONTROLLED BY	ESTROGEN RESPONSE ELEMENTS.			
RP	ESTROGEN RESPONSE ELEMENTS.	"- SUBUNIT: Homodimer (Potential).			

```

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=O75454-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=O75454-2; Sequence=VSP_003702;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE BRAIN, LUNG, BONE MARROW,
CC ADRENAL GLAND, TRACHEA, SPINAL CORD AND THYROID GLAND.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN FETAL BRAIN AND
CC ALSO IN THE FETAL KIDNEY, LUNG AND LIVER.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family, NR3
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF094518; AAC99410.1; -
DR EMBL; AB020639; BAA74955.1; -
DR EMBL; AF058291; AAC39899.1; -
DR EMBL; AF117254; AAD48369.1; -
DR PDB; 1KV6; 25-JAN-03.
DR TRANSPAC; T02766; -
DR TRANSPAC; T04644; -
DR TRANSPAC; T04645; -
DR TRANSPAC; T04648; -
DR Genew; HENC:3474; BSRG.
DR MIM; 602969; -
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR000536; Hormone_rec_1ig.
DR InterPro; IPR001723; Sterhmr_receptor.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR001628; Znf_Casteroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHOMONER.
DR PRINTS; PR0047; STROIDPINGER.
DR PRODOM; PD000035; Znf_Casteroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; ZNF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Activator; Alternative splicing; 3D-structure.
KW DNA BIND 128 193 NUCLEAR RECEPTOR-TYPE.
FT FT DNA BIND 128 148
FT FT ZN FING 164 188
FT FT ZN FING 164 188
FT FT VAREPLIC 1 23 Missing (in isoform Short).
FT FT CONFLICT 151 151 /FTId=VSP_003702.
FT FT CONFLICT 155 155 F->S (IN REF. 3).
FT FT CONFLICT 158 158 T->K (IN REF. 3).
FT FT CONFLICT 271 271 G->A (IN REF. 3).
FT FT CONFLICT 271 271 L->C (IN REF. 3).
FT FT CONFLICT 313 313 V->F (IN REF. 3).
FT FT CONFLICT 313 313 V->F (IN REF. 3).
FT FT CONFLICT 458 458 V->VC (IN REF. 3).
SQ SEQUENCE 458 AA; 51305 MW; 63D36CFD37573152 CRC64;
Query Match 100.0%; Score 2388; DB 1; Length 458;
Best Local Similarity 100.0%; Pred. No. 2e-168;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSEVCLPESPSLHYEBELLCRMKNKDRHNSGSGSFKTEPSSPASLTDVYNNHSPG 60
DB 1 MDSEVCLPESPSLHYEBELLCRMKNKDRHNSGSGSFKTEPSSPASLTDVYNNHSPG 60

```

```

QY 61 SSDASGSYSTWNGHQNGLDSPPLYPADLLGGSGPRKLYDDCSSTIYEDPQTKCEYML 120
DB 61 SSDASGSYSTWNGHQNGLDSPPLYPADLLGGSGPRKLYDDCSSTIYEDPQTKCEYML 120
QY 121 NSMPKRLCYCGDIAGHYGVASCAACRAFFKRTTQGNIEVSCPNTNCEIKRRKSC 180
DB 121 NSMPKRLCYCGDIAGHYGVASCAACRAFFKRTTQGNIEVSCPNTNCEIKRRKSC 180
QY 181 QACRFKCLKVGMIXGVRLDRVRGGRQKRRITDAENSPLYNPOLVQPAKKPYNKLVSH 240
DB 181 QACRFKCLKVGMIXGVRLDRVRGGRQKRRITDAENSPLYNPOLVQPAKKPYNKLVSH 240
QY 241 ILVAEPEKLYAMPDPVTPSDIKALFTLGLADRELVIITGMAKHIFGFTSLIADQMSL 300
DB 241 ILVAEPEKLYAMPDPVTPSDIKALFTLGLADRELVIITGMAKHIFGFTSLIADQMSL 300
QY 301 LQAGAMEILLGVYRSLSPEDLVYADYVIMEDQSKLGLDINNALLQVKKYSMK 360
DB 301 LQAGAMEILLGVYRSLSPEDLVYADYVIMEDQSKLGLDINNALLQVKKYSMK 360
QY 361 LEKEEFTYLKALIALANDSMHIEDVEAVQKLOVLAHALQDYEAGQMEDERRAKMLMT 420
DB 361 LEKEEFTYLKALIALANDSMHIEDVEAVQKLOVLAHALQDYEAGQMEDERRAKMLMT 420
QY 421 LPLRLQSTKAVQHFYNIKLEGRVPMKLFLEMLIAKV 458
DB 421 LPLRLQSTKAVQHFYNIKLEGRVPMKLFLEMLIAKV 458

```

RESULT 2

ERR2 HUMAN STANDARD; PRT; 500 AA.

```

ID ERR2 HUMAN STANDARD; PRT; 500 AA.
AC 095718; O9HCB4;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Steroid hormone receptor ERR2 (Estrogen-related receptor, beta)
DE (ERR-beta) (Estrogen receptor-like 2) (ERR beta-2).
GN ESRRB OR NR3B2 OR ESRB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP Tissue=Testis;
RC MEDLINE=99173874; PubMed=10072763;
RX Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,
RA Liu Q., Caskey C.T.,
RA "Identification of two hERR2-related novel nuclear receptors utilizing
RT bioinformatics and inverse PCR."
RU Gene 228:101-109(1999).
RN [2]
RP SEQUENCE OF 264-500 FROM N.A.
RA Rowen L., Madan A., Qin S., Baradaran L., Birditt B., Bloom S.,
RA Dore M., Dickhoff R., Fleetwood P., Harrison G., Kaur A., Madan A.,
RA Nesbitt R., Traicoiff R., Hood L.,
RA "Sequencing of human chromosome 14q24.3 region."
RU Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family, NR3
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF094517; AAC99409.1; -

```

DR EMBL; AC008050; AAG17472.1; -
 DR HSPSP; P03372; IHCO.
 DR TRANSFAC; T02766; -
 DR Genew; HGNC:3473; ESRRB.
 DR MIM; 602167; -
 DR GO; GO:0004879; F.ligand-dependent nuclear receptor activity; TAS.
 DR InterPro; IPR000536; Hormone_rec_1lg.
 DR InterPro; IPR001723; Steroid_receptor.
 DR InterPro; IPR008946; Str_ncl_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STROHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR Prodom; PD000035; Znf_C4steroid; 1.
 DR SMART; SMO0430; HOL1; 1.
 DR SMART; SMO0399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW zinc-finger.
 FT DNA BIND 103 168 NUCLEAR RECEPTOR-TYPE.
 FT ZN FING 103 123 C4-TYPE.
 FT ZN FING 139 163 C4-TYPE.
 SQ SEQUENCE 500 AA; 55619 MW; C219C84D914DF1C6 CRC64;

Query Match 75.3%; Score 1799; DB 1; Length 500;
 Best Local Similarity 78.1%; Pred. No. 5.5e-125;
 Matches 339; Conservative 50; Mismatches 43; Indels 2; Gaps 2;

QY 24 MSNKRHIDSSGSSFIKTEPSSPASITDSVNHSSPGSSDASGYSSTMGHONGLDSP 83
 DB 1 MSSDRHLGSSCGSFKTEPSSPSGIDALSHHSPGSSDAGGROLAIGHTANGLDSP 60
 QY 84 LYPSPAPILGSGSPVKLYDSCSTIVEDPQTCCEYMLNSMPKRLCTVCGDIASGVHYGYA 143
 DB 61 MFAAGL--GGNPKRSYEDCAGIMEDSAIKCEYMLNAIPKRLCTVCGDIASGVHYGYA 118
 QY 144 SCEACKAFFKRTIQNIIEYSCPAINECEITRRRRSCQACRFMKCLKVGMLKEGYRLDRV 203
 DB 119 SCEACKAFFKRTIQNIIEYSCPAINECEITRRRRSCQACRFMKCLKVGMLKEGYRLDRV 178
 QY 204 RGRGOKYKRRIDANSFYLPQVOPAKKPYNKIVSHLVAEPKITYAMPDPDTPVDSIDK 263
 DB 179 RGRGOKYKRRIDANSFYLPQVOPAKKPYNKIVSHLVAEPDLYAMPDPDTPVDSIDK 238
 QY 264 ALTTICDLADRLVVTIGMAKHIPGPTSTLADQNSLLQSAWMEILLGVYRSLSFDE 323
 DB 239 ALTTICDLADRLVVTIGMAKHIPGPTSTLADQNSLLQSAWMEILLGVYRSLSFDE 298
 QY 324 LYYADDYIMDEDSKLAGLIDNNAILQVKKYKMKLEKFFVTLKALALANSQSHTE 383
 DB 299 LYYADDYIMDEDSKLAGLIDNNAILQVKKYKMKLEKFFVTLKALALANSQSHTE 358
 QY 384 DVEAVQKLDVLEALQDYEAQHEDEPRAGKMLMTPLRLQSTKAVQHFYNTKLESG 443
 DB 359 DVEAVQKLDVLEALQDYEAQHEDEPRAGKMLMTPLRLQSTKAVQHFYNTKLESG 418
 QY 444 VPMKRLPLEMBAK 457
 DB 419 VPMKRLPLEMBAK 432

RESULT 3
 ERR2 RAT
 ID ERR2_RAT STANDARD; PRT; 433 AA.
 AC P11475;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DB Steroid hormone receptor ERR2 (Estrogen-related receptor, beta) (ERR-beta)
 DB beta) (Estrogen receptor-like 2).
 GN ESRRB OR NR382 OR ESRL2 OR ERR2.
 OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=101116;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RA MEDLINE=91173874; PubMed=3267207;
 RA Giguere V., Yang N., Segui P., Evans R.M.;
 RT "Identification of a new class of steroid hormone receptors.";
 RL Nature 331:91-94(1988).
 RP SHOWS THAT REF.1 ORIGINATES FROM RAT.
 RX MEDLINE=99173874; PubMed=10072763;
 RA Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,
 RA Liu Q., Caskey C.T.;
 RT "Identification of two hERR2-related novel nuclear receptors utilizing
 RT bioinformatics and inverse PCR.";
 RL Gene 228:101-109(1999).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
 CC subfamily.
 CC -1- CAUTION: Was originally (Ref.1) thought to originate from human
 CC but was later shown (Ref.2) to be derived from rat.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL; X51417; CAA35779.1; -
 DR HSPSP; P03372; IHCO.
 DR InterPro; IPR000536; Hormone_rec_1lg.
 DR InterPro; IPR001723; Steroid_receptor.
 DR InterPro; IPR008946; Str_ncl_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STROHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR Prodom; PD000035; Znf_C4steroid; 1.
 DR SMART; SMO0430; HOL1; 1.
 DR SMART; SMO0399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW zinc-finger.
 FT DNA BIND 103 168 NUCLEAR RECEPTOR-TYPE.
 FT ZN FING 103 123 C4-TYPE.
 FT ZN FING 139 163 C4-TYPE.
 SQ SEQUENCE 433 AA; 48287 MW; EECC4C5B2E9A3E13 CRC64;

Query Match 74.7%; Score 1784; DB 1; Length 433;
 Best Local Similarity 76.6%; Pred. No. 5.7e-124;
 Matches 333; Conservative 51; Mismatches 49; Indels 2; Gaps 1;

QY 24 MSNKRHIDSSGSSFIKTEPSSPASITDSVNHSSPGSSDASGYSSTMGHONGLDSP 83
 DB 1 MSSDRHLGSSCGSFKTEPSSPSGIDALSHHSPGSSDAGGROLAIGHTANGLDSP 60
 QY 84 LYPSPAPILGSGSPVKLYDSCSTIVEDPQTCCEYMLNSMPKRLCTVCGDIASGVHYGYA 143
 DB 61 MFAAGL--GGNPKRSYEDCAGIMEDSAIKCEYMLNAIPKRLCTVCGDIASGVHYGYA 118
 QY 144 SCEACKAFFKRTIQNIIEYSCPAINECEITRRRRSCQACRFMKCLKVGMLKEGYRLDRV 203
 DB 119 SCEACKAFFKRTIQNIIEYSCPAINECEITRRRRSCQACRFMKCLKVGMLKEGYRLDRV 178
 QY 204 RGRGOKYKRRIDANSFYLPQVOPAKKPYNKIVSHLVAEPKITYAMPDPDTPVDSIDK 263
 DB 179 RGRGOKYKRRIDANSFYLPQVOPAKKPYNKIVSHLVAEPDLYAMPDPDTPVDSIDK 238

RP FUNCTION.
RP MEDLINE=97415618; PubMed=9271417;
RA Sladek R., Bader J.-A., Giguere V.;
RT "The orphan nuclear receptor estrogen-related receptor alpha is a
RT transcriptional regulator of the human medium-chain acyl coenzyme A
RT dehydrogenase gene.";
RL Mol. Cell. Biol. 17:5400-5409(1997).
CC -1- FUNCTION: BINDS TO AN ERR-ALPHA RESPONSE ELEMENT (ERRE) CONTAINING
CC A SINGLE CONSENSUS HALF-SITE, 5'-TNAAGTCA-3'. CAN BIND TO THE
CC MEDIUM-CHAIN ACYL COENZYME A DEHYDROGENASE (MCAD) RESPONSE ELEMENT
CC NRE-1 AND MAY ACT AS AN IMPORTANT REGULATOR OF MCAD PROMOTER.
CC BINDS TO THE C1 REGION OF THE LACTOFERRIN GENE PROMOTER.
CC -1- SUBUNIT: Binds DNA as a monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X51416; CAA35778.1; -;
CC EMBL: L38487; AAB17015.1; -;
CC PIR: A29345; A29345.
CC HSSP: P03372; IHCO.
CC TRANSFAC: T02765; -;
CC Genew: HGNC:3471; ESRR.A.
CC MIM: 601998; -;
CC DR GO: 0005634; C:nucleus; TAS.
CC DR GO: 0003677; F:DNA binding; TAS.
CC DR GO: 0004879; F:ligand-dependent nuclear receptor activity; TAS.
CC DR InterPro: IPR000536; Hormone_rec_1lg.
CC DR InterPro: IPR001723; Strnrm_receptor.
CC DR InterPro: IPR006946; Str_ncl_receptor.
CC DR InterPro: IPR001628; Znf_C4steroid.
CC DR Pfam: PF00104; hormone_rec; 1.
CC DR Pfam: PF00105; zf-C4; 1.
CC DR PRINTS: PR00398; STRDOROMNER.
CC DR PRINTS: PR00047; STRDORFINGER.
CC DR Prodom: PD000035; Znf_C4steroid; 1.
CC DR SMART: SM00430; HOLT; 1.
CC DR SMART: SM00399; Znf_C4; 1.
CC DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
CC KM Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger.
CC FT DNA_BIND 176 241 NUCLEAR RECEPTOR-TYPE.
CC FT ZN_FING 176 196 C4-TYPE.
CC FT ZN_FING 212 236 C4-TYPE.
CC FT CONFLICT 39 44 ATSSRR -> SYKQPA (IN REF. 2).
CC FT CONFLICT 288 288 A -> AA (IN REF. 1).
CC FT CONFLICT 441 450 AEAVEQLEEA -> ERLWSSCKEL (IN REF. 1).
CC SEQUENCE 519 AA; 55439 MW; FA30F3EC7BE7B77FE CRC64;
CC -----
CC Query Match 52.7%; Score 1258.5; DB 1; Length 519;
CC Best Local Similarity 55.1%; Pred. No. 3.7e-85;
CC Matches 256; Conservative 67; Mismatches 79; Indels 63; Gaps 8;
CC -----
CC QY 22 CRRNSNDRHIDSSCS-----FKTEPSSPASLTDVSNHNSPGGSSP-----A 64
CC DB 85 CPTALEPQVTSMSQVVGIEPLTKAEASP-----DSPKSSSETEPEPVAA 135
CC QY 65 SG-SYSTMMGNHNGDPLPSAPILGGSGFVRKLYDDCSSTIYEDPTQCEYMMNM 123
CC DB 136 PGAPAPRCIPGKHEBEDG-----GAGP-----GGGGGKLVLSL 171
CC QY 124 PKRLICVCGDIASGYHGVASCAACAFAFRRTIGNTEYSCATNCEITRRKSCQAC 183
CC DB 172 PKRLICVCGDVASGYHGVASCAACAFAFRRTIGSIEYSCPASNCEITRRKXKQAC 231

QY 184 RFMKCIKVGMLKEGVGLDRVRGGRQYKRRIDENSPYLNPOLVQ-----AKRPN 235
DB 232 RFTKCLRVRGMLKEGVGLDRVRGGRQYKRRPEVDPLPFPPFAGPLAVAGGPRKTA 291
QY 236 KIYSHLLVAPEKTYAMPDPVDSQIKALTTLCIDLADRLVYITGAKHIPPSTSLA 295
DB 292 ALVSHLLVAPEKTYAMPDPVDSQIKALTTLCIDLADRLVYITGAKHIPPSTSLA 351
QY 296 DQMSLLQSAAMEILLIGVYRSISFDELEVADYIMDEQSKLAGLIDNNALIQYVK 355
DB 352 DQMSVLIQSVMEVTLVGVAGRSPLQDELAFADVDIVDEGAAAGGELGALLIQYVR 411
QY 356 YKSKYKKEHEFTLKALIALANSDSMHIIEVAVQKLDVLEALQDYEAQQ---HKEDP 411
DB 412 IQALRLERERERYVLLKALALANSDSVHIIEVAVQKLEALHEALIEYEAAGPGGAER 471
QY 412 RRAGKMLMTLPILKOTSTKAVQHPNVIKEGKPMKLFLEM 456
DB 472 RRAGRLTLTLPILKOTSTKAVLAHFVGLKGVPMKLFLEM 516

RESULT 6
ERR1_MOUSE STANDARD; PRT; 462 AA.
AC 008580;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Steroid hormone receptor ERR1 (Estrogen-related receptor, alpha)
DE (ERR-alpha) (Estrogen receptor-like 1) (Fragment).
OS ESRR.A OR NR3B1 OR ERR1 OR ESTRA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=97415618; PubMed=9271417;
RA Sladek R., Bader J.-A., Giguere V.;
RT "The orphan nuclear receptor estrogen-related receptor alpha is a
RT transcriptional regulator of the human medium-chain acyl coenzyme A
RT dehydrogenase gene.";
RL Mol. Cell. Biol. 17:5400-5409(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Kidney;
RX MEDLINE=98121983; PubMed=9460651;
RA Shigeta H., Zuo W., Yang N., DiAugustine R., Teng C.T.;
RT "The mouse estrogen receptor-related orphan receptor alpha 1:
RT molecular cloning and estrogen responsiveness.";
RL J. Mol. Endocrinol. 19:299-309(1997).
CC -1- FUNCTION: BINDS TO AN ERR-ALPHA RESPONSE ELEMENT (ERRE) CONTAINING
CC A SINGLE CONSENSUS HALF-SITE, 5'-TNAAGTCA-3'. CAN BIND TO THE
CC MEDIUM-CHAIN ACYL COENZYME A DEHYDROGENASE (MCAD) RESPONSE ELEMENT
CC NRE-1 AND MAY ACT AS AN IMPORTANT REGULATOR OF MCAD PROMOTER. MAY
CC FUNCTION AS A MODULATOR OF THE ESTROGEN SIGNALING PATHWAY IN THE
CC UTERUS.
CC -1- SUBUNIT: Binds DNA as a monomer (Probable).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: MOST HIGHLY EXPRESSED IN KIDNEY, HEART, AND
CC BROWN ADIPOCYTES. ALSO FOUND IN UTERUS, CERVIX AND VAGINA.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN AN ORGAN SPECIFIC MANNER THROUGH
CC MID-TO LATE EMBRYONIC DEVELOPMENT WITH PERSISTENT HIGH-LEVEL
CC EXPRESSION IN BROWN ADIPOSE TISSUE AND INTESTINAL MUCOSA.
CC -1- INDUCTION: ACTIVATED BY DIETHYLSILBESTROL (DES) AND ESTRADIOL IN
CC THE UTERUS.
CC -1- PTM: Phosphorylated (Probable).
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 zinc-finger; Steroid-binding; Phosphorylation.
 FT DOMAIN 1 189
 FT DNA BIND 190 255
 FT ZN_FING 190 210
 FT ZN_FING 226 250
 FT DOMAIN 256 315
 FT DOMAIN 316 600
 FT DOMAIN 64 71
 FT MOD_RES 171 174
 FT MOD_RES 109 109
 FT MOD_RES 111 111
 FT MOD_RES 123 123
 FT MOD_RES 172 172
 FT MOD_RES 542 542
 FT CONFLICT 488 488
 SQ SEQUENCE 600 AA; 67030 MW; C9C7D8CACE0F57D8 CRC64;
 Query Match 30.0%; Score 716.5; DB 1; Length 600;
 Best Local Similarity 39.2%; Pred. No. 3.7e-45;
 Matches 168; Conservative 82; Mismatches 128; Indels 51; Gaps 13;
 QY 73 NGHONGLDSPDY---PSAPILGSGGP---VRKLYDCCSTIVDPQTKCE---YMLN 121
 Db 129 HGHQ-----VPEYLENEPSAVAVRTGPAPYRNSDNBRONGRRLSSSEKGMWES 183
 QY 122 SMPRLCLVCGDIASGHYVASCEACAFKKRTIQNEIYSCPATNECEITKRRKSCQ 181
 Db 184 AKETRYCAVCDYASGHYVWASGEGCAFFKRSIQGNDYMCAPATICTIDKRRKSCQ 243
 QY 182 ACFEMKCLVAGLKEGVLEDRVGRGQ-KYKR-RIDAN-----SPYLNQ 226
 Db 244 ACRIRKCTVEMMGGLRKDR-RGRMKIKRORDLEGRENMGSGMRAANLMPSP 302
 QY 227 VQPAKK-----PYNKIVSHLVAEPKITYAMPDPVPBDDIAYLTICDLARELYVI 279
 Db 303 IKHTKKSPPALSTADQVVALDAEPPLITSEVDPSPSEASMMGLTNLADELIVM 362
 QY 280 IGAKKHLPSTSLADQMSLQSAWMEILLGVYRSISFEDELVADYIMDEDSK- 338
 Db 363 INMAKRVPGFDLNIHQVHLLECAMEILMIGVWMSMEKPLLPANILLDRNGKC 422
 QY 339 IAGLIDNNAILQVKKYKSKLEKEKFEVTLKALAN-----SDSMHIEDVEAVOK 390
 Db 423 VEGWVEIFDMLATSSFRMMNLQGEFVCLKSIIILNSGVYTPLSLTKSLKEDIHR 482
 QY 391 LDVLEHALODY--EAGQHMEDP--RRAGKMIMTLPILRQSTKAVCHFNVIKLEKVP 447
 Db 483 VLKINDTLIHMAKAGLTLQOQHRRLAQILLISITRHSNKGMEHLYNMCKKNVPELY 542
 QY 448 KLELEMLEA 456
 Db 543 DLIEMMDA 551
 RESULT 8
 ESRI_CHICK STANDARD; PRT; 589 AA.
 AC P06212;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Estrogen receptor (ER) (estradiol receptor) (ER-alpha).
 GN ESRI OR NR3A1 OR ESR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE=86247578; PubMed=3755102;
 RA Krust A., Green S., Argos P., Kumar V., Walter P., Bornert J.-M.,
 RA Chabon P.;
 RT "The chicken oestrogen receptor sequence: homology with v-erbA and
 RL the human oestrogen and glucocorticoid receptors.";
 CC EMBO J. 5:891-897(1986).
 CC -!- FUNCTION: The steroid hormones and their receptors are involved in
 CC the regulation of eukaryotic gene expression and affect cellular
 CC proliferation and differentiation in target tissues.
 CC -!- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with ER-
 CC beta.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
 CC a DNA-binding domain and a C-terminal steroid-binding domain.
 CC -!- MISCELLANEOUS: In the absence of ligand, steroid hormone receptors
 CC are thought to be weakly associated with nuclear components;
 CC hormone binding greatly increases receptor affinity. The hormone-
 CC receptor complex appears to recognize discrete DNA sequences
 CC upstream of transcriptional start sites.
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X03805; CAA27433.1; -
 CC PIR; A40914; QRCHE.
 CC HSSP; P03372; IHCP.
 CC TRANSFAC; T00264; -
 CC GO; GO:0016585; C:chromatin remodeling complex; ISS.
 CC GO; GO:0016020; C:membrane; ISS.
 CC GO; GO:0030284; F:estrogen receptor activity; ISS.
 CC GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.
 CC GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.
 CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
 CC InterPro; IPR000536; Hormone_rec_1lg.
 CC InterPro; IPR001292; Oestrgn_receptor.
 CC InterPro; IPR001723; Sterhmn_receptor.
 CC InterPro; IPR008946; Str_ncl_receptor.
 CC InterPro; IPR001628; Znf_C4steroid.
 CC Pfam; PF00104; hormone_rec_1.
 CC Pfam; PF02159; Oest_recep_1.
 CC Pfam; PF00105; Zf-C4; 1.
 CC PRINTS; PR00398; STRDHORMNER.
 CC PRINTS; PR00047; STROIDFINGER.
 CC ProDom; PD000035; Znf_C4steroid; 1.
 CC SMART; SM00430; HOL1_1.
 CC SMART; SM00399; Znf_C4; 1.
 CC PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 FT zinc-finger; Steroid-binding.
 FT DOMAIN 1 178
 FT DNA BIND 179 244
 FT ZN_FING 179 199
 FT ZN_FING 215 239
 FT DOMAIN 245 304
 FT DOMAIN 305 589
 SQ SEQUENCE 589 AA; 66746 MW; 1B09223C770A54B CRC64;
 Query Match 29.9%; Score 715; DB 1; Length 589;
 Best Local Similarity 36.2%; Pred. No. 4.7e-45;
 Matches 167; Conservative 90; Mismatches 138; Indels 66; Gaps 12;
 QY 38 FIKTIPSPPASTLDGNNHNS---PGSSSDASGYSYTMNGHONGLDSPDY--PSAPILG 92
 Db 104 FICQAP-----QISPFTHHSQGVVPEYLENEGSGF-----GMREA-APPAPYRPS- 148

```

QY 93 GSGEVRKLYDDCSSSTVEDPQTCCE---YVLNMPKRLCLVCGDILASGTHYVVASCEAC 148
D 149 -----DNRHRSIRERMSSTNEKSGLSMESESTCEYCAVNDYAGYHGWSCGCG 199
QY 149 KAFKRTIQGNIIEVSCPATNECEITKRRKSGCOACRFMKCLKVGMLKEGVLDRVGRQ 208
D 200 KAFKRSRISIOGHNDYWCAPINOCITDKNRKSGCOACRLAKCYEVMKSGIRDRRGEMM 259
QY 209 KYKRRIDAENSP-----YLNPOIVQPAKK-----PYNKIVSHLVAEPE 247
D 260 KOKRRESDSGNSNGASSTELAPLITSPVLVKNKKNKSPALSTLACGMVALLEAEP 319
QY 248 KIYAMPDPVPSDIKALTTLCDLADRELYVLIIGAKHHPGSTSLADQMSLISAWME 307
D 320 IYSEYEDPFRPNEASMTLITNLADRELYVHINAKRYPGVVDLTHDQVHLECAMLE 379
QY 308 ILLIGVYRSLSEFEDLYVADDDYIMEDOSK-LAGLLDINNA1LOLVKKYKSMLEKEEF 366
D 380 ILLIGVYRSLSEFEDLYVADDDYIMEDOSK-LAGLLDINNA1LOLVKKYKSMLEKEEF 439
QY 367 VTLKALALAN-----SDSMHIEDVEAVQKLDVLEALQDYERAGQM---EDPRRAG 415
D 440 VCLKSLILLNSGVYPLSLSTLKSLEERDYIHRVLDKITDTLLHLMKSGSLSQOCHRRUA 499
QY 416 KMLMTPLPLRQTSYKAVQHFYNIKLEGKYPMKHLEFMLEA 456
D 500 QLLILISHIRHMSNKGMEHLYNNCKKNVPLDYLLLEMDA 540
RESULT 9
AC 029040;
ID ESRI_PIG STANDARD; PRT; 595 AA.
RT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).
ESR1 OR NR3A1 OR ESR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
RN [1]
RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=Uterus;
RX MEDLINE=95080454; PubMed=7988744;
RA Beckerkamp D., Jungblut P.W., Thole H.H.;
RT "The C-terminal half of the porcine estradiol receptor contains no
RT post-translational modification: determination of the primary
RT structure."
RT Mol. Cell. Endocrinol. 104:163-172(1994).
CC -1- FUNCTION: Nuclear hormone receptor. The steroid hormones and their
CC receptors are involved in the regulation of eukaryotic gene
CC expression and affect cellular proliferation and differentiation
CC in target tissues (By similarity).
CC -1- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with
CC ESR2. Interacts with NCOA3, NCOA5 and NCOA6 coactivators, leading
CC to a strong increase of transcription of target genes (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC a DNA-binding domain and a C-terminal steroid-binding domain.
CC -1- PTM: Phosphorylated by cyclin A/CDK2 (By similarity).
CC -1- MISCELLANEOUS: In the absence of ligand, steroid hormone receptors
CC are thought to be weakly associated with nuclear components;
CC hormone binding greatly increases receptor affinity. The
CC hormone-receptor complex appears to recognize discrete DNA
CC sequences upstream of transcriptional start sites.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 237167; CAA85524.1; -.
CC PIR: I47140; I47140.
CC HSSP: P03372; IHCP.
DR GO: GO:0016585; C:chromatin remodeling complex; ISS.
DR GO: GO:0016020; C:membrane; ISS.
DR GO: GO:0030284; F:estrogen receptor activity; ISS.
DR GO: GO:0030235; F:nitric-oxide synthase regulator activity; ISS.
DR GO: GO:0016049; P:cell growth; ISS.
DR GO: GO:0030520; P:estrogen receptor signaling pathway; ISS.
DR GO: GO:0045839; P:negative regulation of mitosis; ISS.
DR GO: GO:006355; P:regulation of transcription, DNA-dependent; ISS.
DR InterPro: IPR005366; Hormone_rec_lig.
DR InterPro: IPR001292; Oestrgn_receptor.
DR InterPro: IPR001723; Sterhnm_receptor.
DR InterPro: IPR008946; Str_ncl_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF02159; Oest_reccep_1.
DR Pfam: PF00105; zt-C4_1.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; Znf_C4steroid_1.
DR SMART: SM00430; HOL1_1.
DR SMART: SM00399; Znf_C4_1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KM Zinc-finger; Steroid-binding; Phosphorylation.
FT DOMAIN 1 184 NUCLEAR RECEPTOR-TYPE.
FT DNA_BIND 185 250 C4-TYPE.
FT ZN_FING 185 205 C4-TYPE.
FT ZN_FING 221 245 HINGE.
FT DOMAIN 251 310 STEROID-BINDING.
FT DOMAIN 311 551 POLY-ALA.
FT DOMAIN 64 70 PHOSPHORYLATION (BY CDK2) (BY
FT MOD_RES 104 104 SIMILARITY).
FT MOD_RES 106 106 PHOSPHORYLATION (BY CDK2) (BY
FT MOD_RES 118 118 SIMILARITY).
FT MOD_RES 167 167 PHOSPHORYLATION (BY CK2) (BY SIMILARITY).
FT MOD_RES 537 537 PHOSPHORYLATION (BY TYR-KINASES) (BY
FT SEQUENCE 595 AA; 66361 MW; DEC23DC3085BF001 CRC64;
Query Match 29.9%; Score 713.5; DB 1; Length 595;
Best Local Similarity 42.3%; Pred. No. 6,1e-45;
Matches 159; Conservative 69; Mismatches 105; Indels 43; Gaps 10;
QY 119 MNSMPKRLCLVCGDILASGTHYVVASCEAKAFKRTIQGNIIEVSCPATNECEITKRRK 178
D 176 MESAKERTYCAVNDYAGYHGWSCGCAFFKRSIOGHNDYWCAPINOCITDKNRK 235
QY 179 SCQACRFMKCLKVGMLKEGVLDRVGRQ-KYKR-RIDAEN-----SPYLN 223
D 236 SCQACRLAKCYEVMKSGIRDRRGEMM-KRDR-RGGRMLKHKRQDRDGRNEAVPPGDMRSANLWPS 294
QY 224 POLVQPAKK-----PYNKIVSHLVAEPEKIYAMPDPVPSDIKALTTLCDLADREL 276
D 295 PLIKIKTKNSPVLSTLADQMSILALEAEPPIYSEYDPTRLSEASWMLTNLADREL 354
QY 277 VVILGMACHIGFSTLSLADQMSILQSAWMLILIGVYRSLSEFEDLYVADDDYIMEDQ 336
D 355 VMINIMARKVPEDLSDHDQVHLECAMLEILMIGVYRSLSEFEDLYVADDDYIMEDQ 414
QY 337 SK-LAGLLDINNA1LOLVKKYKSMLEKEEFVTLKALALANS-----DSNH 381
D 415 GKCVGWEIFPMLATLSRFRMMLQGEFFCLKSLILLNSGVYPLSLSTLKSLEEKDH 474

```

```

QY 382 IEDVENQKQDVLEHRLQDYERAGQMEDP-RPAGKMLMTPLRLQSTYKAVQHFYNIKL 440
DB 475 IHRV--LDKTTDTLHLM--KAGLTQQOQHRRLAQLLILSHRHMNSKMGHELYNMKC 530
QY 441 EGVEMKFLFLEMLEA 456
DB 531 KNYVPEYDLLEMLDA 546

RESULT 10
ESR1_POEGU
ID ESR1_POEGU STANDARD; PRT; 587 AA.
AC 091250;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).
GN ESR1 OR NR3A1 OR ESR.
OS Psephila guttata (Zebra finch) (Taeniopygia guttata).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Estrildidae;
OC Estrildinae; Taeniopygia.
OX NCBI_Taxid=59729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97163508; PubMed=9010328;
RA Jacobs E.C., Arnold A.P., Campagnoni A.T.;
RT "Zebra finch estrogen receptor cDNA: cloning and mRNA expression.";
RL J. Steroid Biochem. Mol. Biol. 59:135-145(1996).
CC -1- FUNCTION: The steroid hormones and their receptors are involved in
CC the regulation of eukaryotic gene expression and affect cellular
CC proliferation and differentiation in target tissues.
CC -1- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with ER-
CC beta (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC a DNA-binding domain and a C-terminal steroid-binding domain.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L79911; AAB81108.1; -.
CC HSSP: P03372; IHCP.
DR GO: GO:0016585; C:chromatin remodeling complex; ISS.
DR GO: GO:0016020; C:membrane; ISS.
DR GO: GO:0030284; P:estrogen receptor activity; ISS.
DR GO: GO:0030235; P:nitric-oxide synthase regulator activity; ISS.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
DR InterPro: IPR005362; Hormone_rec_1lg.
DR InterPro: IPR001292; Oestrgn_receptor.
DR InterPro: IPR001723; Strhnm_receptor.
DR InterPro: IPR008946; Str_ncl_receptor.
DR InterPro: IPR001628; Znf_Csteroi.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF02159; Oest_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; Znf_Csteroi.1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;

```

```

KW Zinc-finger; Steroid-binding.
FT DOMAIN 1 176 MODULATING.
FT DNA_BIND 177 242 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 177 197 C4-TYPE.
FT ZN_FING 213 237 C4-TYPE.
FT DOMAIN 243 302 HINGE.
FT DOMAIN 303 587 STEROID-BINDING.
SQ SEQUENCE 587 AA; 66553 MW; 2B254168A7A910AB CRC64;

Query Match 29.8%; Score 712.5; DB 1; Length 587;
Best Local Similarity 36.1%; Pred. No. 7.1e-45;
Matches 171; Conservative 90; Mismatches 144; Indels 69; Gaps 14;

QY 33 SSCSSP--IKTEPSSPASLTDS-----VNHS---PGSSDASGSYSTNMGHQNGLD 80
DB 84 SSLAGFHSLSNVSPSPVVFQLQAPHMSPFHHNSQGVPLYLNDQGSF-----GMRA-- 136
QY 81 SPLLFSAPPLGSGGVKRYDDCSSTIYEDPQTKC---YVNSMPKRLCIVCGDIAS 136
DB 137 APPAF-----YRPNSDNRHSIRERMSANEKGSLSVESTKETEYCAVCDYAS 185
QY 137 GYHYVASCEACAKAFKRTIOGNIEXSCPATNCEITKRRRSQCAQREFKCLKYGMKE 196
DB 186 GHYGVMSCEGCACAFKRSIQGHNDYMCPTNQCITDKNRKSCQACRLKCYEVEGKMG 245
QY 197 GVALDRVRCGR-QKYRRRI DAENSP-----YLNPLYOPAKK-----PY 234
DB 246 GIKRDR-RGGRVWKQRQREEDQSRNGEASTELRAPTIWASPLVYKHNKNSPALSLTA 304
QY 235 NKIVSHLVAAPEKRTYAMDPEVPYPSDITKALTLCILADRELVIYTGAKHIFGFTLSL 234
DB 305 EDVVSALDEAPPLVSEDPENRPFNEASMMTLTLVADRELVHMTNMAKRVGFDVLLT 364
QY 295 ADQMSLLQSAWMEIILGVYRSLSPEDELVYADVIYMDQSK-LAGLLDNNAILQIV 353
DB 365 HQVHLLECAWLEILMTIGVYMSMEHPGKLPENLLDNRQKCYEGAWEIFDMLLATA 424
QY 354 KYKSMKLEKEBEVTLKATALAN-----SDSMIEPVEAVQKQDVLEHRLQDYERAG 405
DB 425 ARFRMNLQGEFEVCKSLILNSGVYFTLSLKSLEKEDYHHRVLDKTTDTLHLMMAK 484
QY 406 QHM--EDPRRAGKMLMTPLRLQSTYKAVQHFYNIKLKGVPKMLFLEMLEA 456
DB 485 SGLSLQOQHRRLAQLLILSHRHMNSKMGHELYNMCKKNVPLDLEMLDA 538

RESULT 11
ESR1_MESAU
ID ESR1_MESAU STANDARD; PRT; 595 AA.
AC Q9QZD5; Q9QZG6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).
GN ESR1 OR NR3A1 OR ESR.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_Taxid=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=20197937; PubMed=10731637;
RA Bhat H.K., Vagstad J.V.;
RT "Hamster estrogen receptor cDNA: cloning and mRNA expression.";
RL J. Steroid Biochem. Mol. Biol. 72:47-53(2000).
RN [2]
RP SEQUENCE OF 98-291 FROM N.A.
RA Jones J.E., Carpenter C.D., Imbers L.S., Petersen S.L., Wade G.N.;
RT "Return of lordosis after food deprivation and refeeding in Syrian
KW hamsters." (SEP-1999) to the EMBL/GenBank/DBJ databases.

```

	-1- FUNCTION:	Nuclear hormone receptor. The steroid hormones and their receptors are involved in the regulation of eukaryotic gene expression and affect cellular proliferation and differentiation in target tissues (By similarity).
CC	-1- SUBUNIT:	Binds DNA as a homodimer. Can form a heterodimer with ESR2. Interacts with NCOA3, NCOA5 and NCOA6 coactivators, leading to a strong increase of transcription of target genes (By similarity).
CC	-1- SUBCELLULAR LOCATION:	Nuclear.
CC	-1- DOMAIN:	Composed of three domains: a modulating N-terminal domain, a DNA-binding domain and a C-terminal steroid-binding domain.
CC	-1- SIMILARITY:	Belongs to the nuclear hormone receptor family. NR3 subfamily.
CC		-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_licenses@isb-sib.ch).	
CC		-----
DR	EMBL;	AF181077; AAD53956.1; -
DR	EMBL;	AF184276; AAF02721.1; -
DR	HSSP;	P03372; IHCO.
DR	GO;	GO:0016585; C:chromatin remodeling complex; ISS.
DR	GO;	GO:0016020; C:membrane; ISS.
DR	GO;	GO:0030284; F:estrogen receptor activity; ISS.
DR	GO;	GO:0030285; F:nitric-oxide synthase regulator activity; ISS.
DR	GO;	GO:0016049; P:cell growth; ISS.
DR	GO;	GO:0030520; P:estrogen receptor signaling pathway; ISS.
DR	GO;	GO:0045839; P:negative regulation of mitosis; ISS.
DR	InterPro;	IPR000536; Hormone_rec_1lg.
DR	InterPro;	IPR001292; Oestrin receptor.
DR	InterPro;	IPR001723; Sterdmn receptor.
DR	InterPro;	IPR008946; Srl_ncl_receptor.
DR	InterPro;	IPR001528; Znf_Casteroid.
DR	Pfam;	PF00104; hormone_rec; 1.
DR	Pfam;	PF02159; Oest_recip; 1.
DR	Pfam;	PF00105; zf-C4; 1.
DR	PRINTS;	PR00398; STRODHOMONER.
DR	PRINTS;	PR00047; STROIDINGER.
DR	ProDom;	PD000035; Znf_Casteroid; 1.
DR	SMART;	SM00430; HOL1; 1.
DR	SMART;	SM00399; Znf_C4; 1.
KW	PROSITE;	PS00031. NUCLEAR RECEPTOR; 1.
KW	Receptor;	Transcription regulation; DNA-binding; Nuclear protein;
KW	Zinc-finger;	Steroid-binding; Phosphorylation.
FT	DOMAIN	1 .. 184 MODULATING.
FT	DNA_BIND	185 .. 250 NUCLEAR RECEPTOR-TYPE.
FT	ZN_FING	185 .. 205 C4-TYPE.
FT	ZN_FING	221 .. 245 C4-TYPE.
FT	DOMAIN	251 .. 310 HINGE.
FT	DOMAIN	311 .. 551 STEROID-BINDING.
FT	DOMAIN	112 .. 115 POLY-PRO.
FT	DOMAIN	166 .. 169 POLY-SER.
FT	MOD_RES	103 PHOSPHORYLATION (BY CDK2) (BY SIMILARITY).
FT	MOD_RES	105 PHOSPHORYLATION (BY CDK2) (BY SIMILARITY).
FT	MOD_RES	118 PHOSPHORYLATION (BY SIMILARITY).
FT	MOD_RES	167 PHOSPHORYLATION (BY CK2) (BY SIMILARITY).
FT	MOD_RES	537 PHOSPHORYLATION (BY TYR-KINASES) (BY SIMILARITY).
FT	CONFLICT	98 OL -> T (IN REF. 2).
FT	CONFLICT	144 S -> T (IN REF. 2).
FT	CONFLICT	199 G -> EVEQI (IN REF. 2).
FT	CONFLICT	290 NL -> KG (IN REF. 2).
SQ	SEQUENCE	595 AA; 66833 MW; E4D803B4FDB257E CRC64;
Query Match	29.8%;	Score 711.5; DB 1; Length 595;
Best Local Similarity	36.4%;	Pred No. 8;6e-45;

	Matches	177;	Conservative	79;	Mismatches	145;	Indels	77;	Gaps	12;		
QY	56	HSFGSGSSDAGGSYS	STTMNGHONG	GLDSP---	LYPSAPILG-----	GS6PRK	LYDDCS	106				
Db	79	YGSQSSKATAGCGNS	SLGF	FPOLNSVSP	PELMLHP	PPPLQ	PSFLPHGQ	QVPLYL	ENESA	138		
QY	107	TIYEDPQTC	OEYMLNS-----			MPKRL	CLVCGD	IASG	YHAG	141		
Db	139	YAYRDSPP	PAFRKNS	DNRRQSG	RERLSS	SSSEK	SMAMES	VKEIR	RYC	AVCDYASG	198	
QY	142	VASGEACKAF	EFKRTIT	QGNIEYS	CPATNECE	ITKRR	RKSCQA	CRMK	CLKVOML	KEGVLD	201	
Db	199	VMSCEGCKAF	EFKRSIQ	GHNDYMC	ATNQCTID	KNRKRS	COA	CR	LRKCYE	GMKG	198	
QY	202	RVRGGRQ	KTKRR	LD-----		ANSTY	INQLVQ	VP	PAKK-----	PYNK	239	
Db	259	R-RGRGRL	KKRR	RDDLEGR	NDMGP	SGDM	RATN	IMP	SPVLK	HKYKNS	PA	317
QY	240	HLVAEPEK	YVAMPDP	TVDP	BDIDKAL	FTL	CLAD	RE	LVII	IGMAKH	IPG	299
Db	318	ALLDAE	EPPLIY	SEVDP	SRP	BSA	MGMLT	MLAD	RELVH	MTNM	KRVG	377
QY	300	LLQSAWE	ELLILGV	YVSLSE	FED	ELVY	ADYIM	EDQSK	-LAG	LID	NNAILQ	358
Db	378	LLECAWE	ELIMIG	ELTMR	SMER	HPG	KLTF	APN	LIL	DRNQCK	CEVG	437
QY	359	MKLEKE	FVTLK	AL	ALANS-----		DSMH	ED	EA	VO	KDVL	404
Db	438	MDLOGEE	FVCL	CKSLI	ILNS	GVYTL	SLT	KSLE	ED	H	IRV-	493
QY	405	QGHMEDP	-RR	AKML	MTLP	PL	TR	OT	STK	VA	GFVNI	456
Db	494	GTLTQ	QCHRL	LAQ	LIL	IL	ISH	RH	NS	KN	KEH	546

RESULT 12

ESRL MOUSE STANDARD, PRT, 599 AA.

ID ESRL MOUSE STANDARD, PRT, 599 AA.

AC P19785; Q9JUT5; Q9GY51; Q9GY52;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Estrone receptor (ER) (Estradiol receptor) (ER-alpha).

GN ESRL OR NR3A1 OR ESR OR ESR1 OR ESR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid:10090;

PN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=uterus;

RX MEDLINE=91042558; PubMed=2484714;

RA White R., Lees J.A., Needham M., Ham J., Parker M.;

RT "Structural organization and expression of the mouse estrogen receptor";

RL Mol. Endocrinol. 1:735-744(1987).

LN [2]

LN SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Bone, and Thymus;

RX MEDLINE=22354683; PubMed=1246651;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

RA Nikaido I., Otsu N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotojori T.,

RA Baldarini L.R., Hill D.F., Bult C., Hume D.A., Quackenbush J.,

RA Schmitt L.M., Knapin A., Matsuda H., Batalov S., Belser K.W.,

RA Balle J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,

RA Datta E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Girimond S., Guslinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,

RA Kongaya A.R., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,

RA Maglott D.R., Maltzins L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Seton M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirokane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs,"
RT Nature 420:563-573 (2002).
RN [3]
RP SEQUENCE OF 1-22 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=20359284; PubMed=10899303;
RA Kos M., O'Brien S., Flourio G., Cannon F.,
RT "Tissue-specific expression of multiple mRNA variants of the mouse
RT estrogen receptor alpha gene,"
RL PNAS Lett. 477:15-20(2000).
RN [4]
RP SEQUENCE OF 269-599 FROM N.A.
RC STRAIN=SJL/J, and B10.S/J; TISSUE=Spleen;
RA Ma R.Z., Teuscher C.,
RT "Screening for candidate genes of mouse autoimmune diseases,"
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP CARBOHYDRATE-LINKAGE SITE SER-575.
RX MEDLINE=97153020; PubMed=8999954;
RA Jiang M.S., Hart G.W.,
RT "A subpopulation of estrogen receptors are modified by O-linked
RT N-acetylglucosamine,"
RL J. Biol. Chem. 272:2421-2428(1997).
RN [6]
RP CARBOHYDRATE-LINKAGE SITES SER-10; THR-50 AND SER-575.
RX MEDLINE=21124487; PubMed=11226831;
RA Cheng X., Hart G.W.,
RT "Glycosylation of the murine estrogen receptor-alpha,"
RL J. Steroid Biochem. Mol. Biol. 75:147-158(2000).
RN [7]
RP INTERACTION WITH NCOA3.
RX MEDLINE=97336097; PubMed=9192892;
RA Torchia J., Rose D.W., Inostroza J., Kamei Y., Westin S., Glass C.K.,
RT "The transcriptional co-activator p/CIP binds CBP and mediates
RT nuclear-receptor function,"
RL Nature 387:677-684(1997).
RN [8]
RP INTERACTION WITH NCOA6.
RX MEDLINE=20250907; PubMed=10788465;
RA Zhu Y.-J., Kan L., Qi C., Kanwar Y.S., Yelandi A.V., Rao M.S.,
RA Reddy J.K.,
RT "Isolation and characterization of peroxisome proliferator-activated
RT receptor (PPAR) interacting protein (PRIP) as a coactivator for
RT PPAR,"
RL J. Biol. Chem. 275:13510-13516(2000).
RN [9]
RP FUNCTION: Nuclear hormone receptor. The steroid hormones and their
RN receptors are involved in the regulation of eukaryotic gene
RN expression and affect cellular proliferation and differentiation
RN in target tissues.
RN [10]
RP SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with
RN ESR2. Interacts with NCOA3 and NCOA6 coactivators, leading to a
RN strong increase of transcription of target genes. Interacts with
RN NCOA5 (by similarity).
RN [11]
RP SUBCELLULAR LOCATION: Nuclear.
RN [12]
RP DOMAIN: Composed of three domains: a modulating N-terminal domain,
RN a DNA-binding domain and a C-terminal steroid-binding domain.
RN [13]
RP PTM: Phosphorylated by cyclin A/CDK2 (by similarity).
RN [14]
RP MISCELLANEOUS: In the absence of ligand, steroid hormone receptors

CC are thought to be weakly associated with nuclear components;
CC hormone binding greatly increases receptor affinity. The
CC hormone-receptor complex appears to recognize discrete DNA
CC sequences upstream of transcriptional start sites.
CC [15]
CC SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M38651; AAA37580.1; -
CC EMBL; AK036627; BAC29510.1; -
CC EMBL; AK041525; BAC30973.1; -
CC EMBL; AJ276597; BAC85618.1; -
CC EMBL; AF128221; AAF22562.1; -
CC EMBL; AF128220; AAF22564.1; -
CC PIR; A40061; QRMSE.
CC HSP; P03372; 1HCP.
CC TRANSFAC; T00259; -.
CC GLYCONSULEDB; P19785; -.
CC MGD; MGI:1352467; Esrl.
CC GO; GO:0016585; Chromatin remodeling complex; ISS.
CC GO; GO:0016020; C:membrane; ISS.
CC GO; GO:0030284; C:membrane; ISS.
CC GO; GO:0030235; F:estrogen receptor activity; ISS.
CC GO; GO:0016049; F:cell growth; NAS.
CC GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.
CC GO; GO:0045839; P:negative regulation of mitosis; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
CC InterPro; IPR000536; Hormone_rec_1lg.
CC InterPro; IPR01292; Oestrgn_receptor.
CC InterPro; IPR001723; Stergm_receptor.
CC InterPro; IPR008946; Str_ncl_receptor.
CC InterPro; IPR001628; Znf_ncl_receptor.
CC Pfam; PF00104; hormone_rec; 1.
CC Pfam; PF02159; Oest_rec; 1.
CC Pfam; PF00105; ZF-C4; 1.
CC PRINTS; PRO0398; STERHORMONER.
CC PRINTS; PRO0047; STROIDPFINGER.
CC ProDom; PD000035; Znf_C4steroid; 1.
CC SMART; SM00430; HOL1; 1.
CC SMART; SM00399; ZNF_C4; 1.
CC PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger; Steroid-binding; Phosphorylation; Glycoprotein.
CC DOMAIN
CC 1 188
CC FT DNA_BIND 189 254
CC FT ZN_FING 189 209
CC FT ZN_FING 225 249
CC FT DOMAIN 255 314
CC FT DOMAIN 315 599
CC FT DOMAIN 64 72
CC FT MOD_RES 108 108
CC FT
CC FT MOD_RES 110 110
CC FT
CC FT MOD_RES 122 122
CC FT MOD_RES 171 171
CC FT MOD_RES 541 541
CC FT
CC FT CARBOHYD 10 10
CC FT CARBOHYD 50 50
CC FT CARBOHYD 575 575
CC FT
CC FT VARIANT 591 591
CC FT CONFLICT 269 269
CC FT
CC SQ SEQUENCE 599 AA; 66955 MW; 05F5E2FC21CC0A8B CRC64;

```

Query Match      29.7%; Score 710; DB 1; Length 599;
Best Local Similarity 41.8%; Pred. No. 1,1e-44;
Matches 157; Conservative 72; Mismatches 104; Indels 43; Gaps 10;

QY 119 MNSMPKRLCLVCGDIAGYGVAVSCGACAFKFRITQGNIEYSCPATNECEITKRRK 178
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 180 MESAKETRYCAVNCNDYASGHHYGVWSCGECCKAFKRSIQGNIDYMCATNQCITDKRRK 239
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY 179 SCQACREMKCLVGMKEGVRLDVRGGRQ-KYKR-RIDAEN-----SPYLN 223
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 240 SCQACGRKRCYGVGMKGGIRKDR-RGGHMLKHKRDRDLDEGNENMGASGDMRAALWPS 298
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 224 POLVOPAKK-----PYMKIVSHLVAREPKYAMPDPVPSDICALTTICDLADREL 276
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 299 PLVIRKTKKNSPALSTADQWVSALDAPPMIYSYDSDSRPSEASMMGLTNLADREL 358
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 277 VVILGWAKHIGPFSLSLADQMSLQSAWMEILLGVVYRSLSFEDELVYADYIMDEQ 336
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 359 VHMIMNAKRVPGGDLNHDQVHLLECAMEILMIGLVWRSMHPKLLFAPNILLDRNQ 418
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 337 SK-LAGLIDLNNAIQLQVKKYKSMKLEKEFVTLKALANS-----DSMH 381
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 419 GKVEGAWVEIFDMILATSSRFMMNIQGEFEVCLKSIILNSGVYTFLSSTLKLEKDH 478
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 382 IEDVEAVOKLQVTLHVALQDYEAQGHMEDP-RPAGKMLMTLPRLRQSTKAVOHFYNIKL 440
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 479 IHRV--LDKTTDPLIHMA--KAGLTLQOHRRLAQLLILSHIRHMSKMGHLYMKC 534
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 441 EGVPMKLFLEMLEA 456
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 535 KNYVPLYLLEMLDA 550
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 13
ESR1 HORSE STANDARD; PRT; 594 AA.
ID ESR1 HORSE STANDARD; PRT; 594 AA.
AC Q9TVG8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Estrogen receptor (ER) (Estriadiol receptor) (ER-alpha).
GN ESR1 OR NR3A1 OR ESR.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA McDowell K.J., Adams M.H., Green M.L., Cleaver B.D., Sharp D.C.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Nuclear hormone receptor. The steroid hormones and their
   receptors are involved in the regulation of eukaryotic gene
   expression and affect cellular proliferation and differentiation
   in target tissues (By similarity).
CC -1- SUBUNIT: Binds DNA as a homodimer. Can form a heterodimer with
   ESR2. Interacts with NCOA3, NCOA5 and NCOA6 coactivators, leading
   to a strong increase of transcription of target genes (By
   similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: Composed of three domains: a modulating N-terminal domain,
   a DNA-binding domain and a C-terminal steroid-binding domain.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
   subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AF124093; AAD17316.1; -.

```

```

DR HSSP; P03372; IHCP.
DR GO; GO:0016585; C:chromatin remodeling complex; ISS.
DR GO; GO:0016020; C:membrane; ISS.
DR GO; GO:0030284; F:estrogen receptor activity; ISS.
DR GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.
DR GO; GO:0016049; F:cell growth; ISS.
DR GO; GO:0030520; P:estrogen receptor signaling pathway; ISS.
DR GO; GO:0045839; P:negative regulation of mitosis; ISS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.
DR InterPro; IPR00536; Hormone rec. 1lg.
DR InterPro; IPR001292; Oestrgn_receptor.
DR InterPro; IPR001723; Sterhma_receptor.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF02159; Oest_recip; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00396; STRDHOMONER.
DR PRODOM; PR00047; STROIDPTNGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; ZNF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
   Zinc-finger; Steroid-binding; Phosphorylation.
KW DOMAIN 1 184
   FT DNA_BIND 185 250
   FT ZN_FING 185 205
   FT ZN_FING 221 245
   FT ZN_FING 251 310
   FT DOMAIN 311 550
   FT DOMAIN 64 70
   FT MOD_RES 104 104
   FT MOD_RES 106 106
   FT MOD_RES 118 118
   FT MOD_RES 167 167
   FT MOD_RES 536 536
   SQ SEQUENCE 594 AA; 66103 MW; DD36CA7C247A895 CRC64;

```

```

Query Match      29.5%; Score 704; DB 1; Length 594;
Best Local Similarity 41.1%; Pred. No. 3,1e-44;
Matches 154; Conservative 72; Mismatches 107; Indels 42; Gaps 9;

QY 119 MNSMPKRLCLVCGDIAGYGVAVSCGACAFKFRITQGNIEYSCPATNECEITKRRK 178
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 176 MESAKETRYCAVNCNDYASGHHYGVWSCGECCKAFKRSIQGNIDYMCATNQCITDKRRK 235
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY 179 SCQACREMKCLVGMKEGVRLDVRGGRQ-KYKR-RIDAEN-----AENSP---YLN 223
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 236 SCQACGRKRCYGVGMKGGIRKDR-RGGHMLKHKRDRDLDEGNENMGASGDMRAALWPS 294
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 224 POLVOPAKK-----PYMKIVSHLVAREPKYAMPDPVPSDICALTTICDLADREL 276
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 295 PLVIRKTKKNSPALSTADQWVSALDAPPMIYSYDSDSRPSEASMMGLTNLADREL 354
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 277 VVILGWAKHIGPFSLSLADQMSLQSAWMEILLGVVYRSLSFEDELVYADYIMDEQ 336
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 355 VHMIMNAKRVPGGDLNHDQVHLLECAMEILMIGLVWRSMHPKLLFAPNILLDRNQ 414
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 337 SK-LAGLIDLNNAIQLQVKKYKSMKLEKEFVTLKALANS-----DSMH 381
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 415 GKVEGAWVEIFDMILATSSRLRMMNIQGEFEVCLKSIILNSGVYTFLSSTLKLEKDH 474
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 382 IEDVEAVOKLQVTLHVALQDYEAQGHMEDP-RPAGKMLMTLPRLRQSTKAVOHFYNIKL 441
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 475 IHRV--LDKMDTDLIHMA--KAGLTLQOHRRLAQLLILSHIRHMSKMGHLYMKC 530
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 442 EGVPMKLFLEMLEA 456
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 531 NNYVPLYLLEMLDA 545
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

RESULT 14
ESR1_HUMAN STANDARD; PRT; 595 AA.
AC P03372; Q13511; Q14276; Q9NU51; Q9UDZ7; Q9UIS7;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Estrogen receptor (ER) (Estradiol receptor) (ER-alpha).
GN ESR1 OR NR3A1 OR ESR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=86122927; PubMed=3753802;
RA Greene G.L., Gilna P., Waterfield M., Baker A., Hort Y., Shine J.;
RT "Sequence and expression of human estrogen receptor complementary
RT DNA.";
RL Science 231:1150-1154(1986).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=86146892; PubMed=3754034;
RA Green S., Walter P., Kumar V., Krust A., Bornert J.-M., Argos P.,
RA Chambon P.;
RT "Human oestrogen receptor cDNA: sequence, expression and homology to
RT v-erb-A.";
RL Nature 320:134-139(1986).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM LONG), AND VARIANT ASP-411 INS.
RC TISSUE=breast;
RX MEDLINE=96174655; PubMed=8600466;
RA Pink J.J., Wu S.Q., Wolf D.W., Bilimoria M.M., Jordan V.C.;
RT "A novel 80 kDa human estrogen receptor containing a duplication of
RT exons 6 and 7.";
RL Nucleic Acids Res. 24:962-969(1996).
RN [4]
RP SEQUENCE OF 152-595 FROM N.A. (ISOFORM LONG).
RX MEDLINE=20084372; PubMed=10619354;
RA Schubert E.L., Lee M.K., Newman B., King M.C.;
RT "Single nucleotide polymorphisms (SNPs) in the estrogen receptor gene
RT and breast cancer susceptibility.";
RL J. Steroid Biochem. Mol. Biol. 71:21-27(1999).
RN [5]
RP SEQUENCE OF 216-434 FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=breast carcinoma;
RX MEDLINE=93153765; PubMed=7916651;
RA Pfeiffer U., Fecarotta E., Castagnetta L., Vidali G.;
RT "Estrogen receptor variant messenger RNA lacking exon 4 in estrogen-
RT responsive human breast cancer cell lines.";
RL Cancer Res. 53:741-743(1993).
RN [6]
RP SEQUENCE OF 110-117, PHOSPHORYLATION, AND MUTAGENESIS.
RX MEDLINE=96026869; PubMed=7476978;
RA Joel P.B., Traish A.M., Lannigan D.A.;
RT "Estradiol and phorbol ester cause phosphorylation of serine 118 in
RT the human estrogen receptor.";
RL Mol. Endocrinol. 9:1041-1052(1995).
RN [7]
RP SEQUENCE OF 354-548 FROM N.A.
RC TISSUE=breast carcinoma;
RA Naundorff H., Becker M., Flabio C., Buettner B., Fichtner I.;
RT "Mechanisms of acquired tamoxifen resistance in a xenotransplanted
RT human breast carcinoma.";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 413-595 FROM N.A.
RA Parker A.;
RP Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE OF 532-542, AND PHOSPHORYLATION.

RX MEDLINE=95280953; PubMed=7539106;
RA Arnold S.F., Obourn J.D., Jaffe H., Notides A.C.;
RT "Phosphorylation of the human estrogen receptor on tyrosine 537 in
RT vivo and by src family tyrosine kinases in vitro.";
RL Mol. Endocrinol. 9:24-33(1995).
RN [10]
RP INTERACTION WITH NCOA6.
RX MEDLINE=20036574; PubMed=10567404;
RA Lee S.-K., Anzick S.L., Choi J.-E., Bubendorf L., Guan X.-Y.,
RA Jung Y.-K., Kallioniemi O.P., Kononen J., Trent J.M., Azorsa D.,
RA Jhun B.-H., Cheong J.H., Lee Y.C., Meltzer P.S., Lee J.W.;
RT "A nuclear factor ASC-2, as a cancer-amplified transcriptional
RT coactivator essential for ligand-dependent transactivation by nuclear
RT receptors in vivo.";
RL J. Biol. Chem. 274:34283-34293(1999).
RN [11]
RP INTERACTION WITH NCOA5.
RX MEDLINE=20565767; PubMed=11113208;
RA Sauve F., McBroome L.D.B., Gallant J., Moraitis A.N., Labrie F.,
RA Giguere V.;
RT "CIA, a novel estrogen receptor coactivator with a bifunctional
RT nuclear receptor interacting determinant.";
RL Mol. Cell. Biol. 21:343-353(2001).
RN [12]
RP PHOSPHORYLATION.
RX MEDLINE=95140025; PubMed=7838153;
RA Arnold S.F., Obourn J.D., Jaffe H., Notides A.C.;
RT "Serine 167 is the major estradiol-induced phosphorylation site on the
RT human estrogen receptor.";
RL Mol. Endocrinol. 8:1208-1214(1994).
RN [13]
RP PHOSPHORYLATION OF SER-104 AND SER-106, AND MUTAGENESIS.
RX MEDLINE=99357754; PubMed=10428798;
RA Rogatsky I., Trowbridge J.M., Garabedian M.J.;
RT "Potential of human estrogen receptor alpha transcriptional
RT activation through phosphorylation of serines 104 and 106 by the
RT cyclin A-CDK2 complex.";
RL J. Biol. Chem. 274:22296-22302(1999).
RN [14]
RP STRUCTURE BY NMR OF 180-262.
RX MEDLINE=91061891; PubMed=2247153;
RA Schwabe J.W.E., Neuhaus D., Rhodes D.;
RT "Solution structure of the DNA-binding domain of the estrogen
RT receptor.";
RL Nature 348:458-461(1990).
RN [15]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 180-262.
RX MEDLINE=94037103; PubMed=8221895;
RA Schwabe J.W.E., Chapman L., Finch J.T., Rhodes D.;
RT "The crystal structure of the estrogen receptor DNA-binding domain
RT bound to DNA: how receptors discriminate between their response
RT elements.";
RL Cell 75:567-578(1993).
RN [16]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 305-548.
RX MEDLINE=97478539; PubMed=9338790;
RA Bizozowski A.M., Pike A.C.W., Dauter Z., Hubbard R.B., Bonn T.,
RA Enzozowski O., Oelman L., Greene G.L., Gustafsson J.-A., Carlquist M.;
RT "Molecular basis of agonism and antagonism in the oestrogen
RT receptor.";
RL Nature 389:753-758(1997).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 306-544.
RX MEDLINE=98263297; PubMed=9600906;
RA Tanenbaum D.M., Wang Y., Williams S.P., Sigler P.B.;
RT "Crystallographic comparison of the estrogen and progesterone
RT receptor's ligand binding domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:5998-6003(1998).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 294-554.
RX MEDLINE=99091051; PubMed=9875847;
RA Shiau A.K., Barstad D., Loria P.M., Cheng L., Kushner P.J.,
RA Agard D.A., Greene G.L.;

DR	InterPro: IPR008946; Str_ncl_receptor.
DR	InterPro: IPR001628; ZnF_C4steroid.
DR	Pfam: PF00104; hormone_rec; 1.
DR	Pfam: PF02159; Oest_recep; 1.
DR	Pfam: PF00105; zf-C4; 1.
DR	PRINTS: PR00398; STRDHORMONER.
DR	PRINTS: PR00047; STROIDPTRNGR.
DR	ProDom: PD000035; ZnF_C4steroid; 1.
DR	SMART; SM00430; HOLT; 1.
DR	SMART; SM00399; ZnF_C4; 1.
DR	PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW	Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW	Zinc-finger; Steroid-binding.
FT	DOMAIN 1 179 MODULATING.
FT	DNA_BIND 180 245 NUCLEAR RECEPTOR-TYPE.
FT	ZN_FING 180 200 C4-TYPE.
FT	ZN_FING 216 240 C4-TYPE.
FT	DOMAIN 246 302 HINGE.
FT	DOMAIN 303 586 STEROID-BINDING.
SQ	SEQUENCE 586 AA; 66080 MW; 0EDC77BBD6F08BF CRC64;

Query Match	29.1%;	Score 694;	DB 1;	Length 586;
Best Local Similarity	34.5%;	Pred NO 1	60.43;	

Matches	161;	Conservative	86;	Mismatches	140;	Indels	80;	Gaps	12;
---------	------	--------------	-----	------------	------	--------	-----	------	-----

QY 43 PSSPA-----SLTDSVNHSS--PGSSSDAAGSY-----SSTMGCHQNGLD 80
 Db 99 PPSPIVETLAKLPOLSPFIHHGQOVYYIESEGTAVREAPPTFYRSSSDNRQSGRE 158
 QY 81 SPPILPSADILGSGGVKLYDDCSSTIYBDPOTKCEYMLNSMPKRLCTVCGDIASGYH 140
 Db 159 R-----MSSANDKGPSS-----MSTMETRYCANCSYASGYH 192
 QY 141 GVASCEACAKPFKRTIOGNIYESCPATNCEITKRRRKSQACRPFMKCLKYMKLKGCVTL 200
 Db 193 GWSSEGGCAFFKRSTIOGHNDYICPAINOCTIDKNRKSQACRLKCYEVMGKGIKRX 252
 QY 201 DRRVGR-----QYKXRIDAEN-----SPYLNPOLVDBAKPY-----NKTYSHL 241
 Db 253 DR-RGRRLTKHRKQKECEQKNDVDPSEIRTASIWNPBYSKMSKSPVLSLTPEQLISAL 311
 QY 242 LVAEBEKIYAMPDPVBPDSIDKALITLCLADRELVIITGMANKHGFSTLSIADOMSLI 301
 Db 312 MEAEPLIYVSEHDSKTPLESEASMTLLNLMLARELVHMIMNAKRVGPFVDLTLDQVHLL 371
 QY 302 QSAMNELILGVYKSLSEFEDLVADYIMDEDSK-LAGLIDLNALLQVKKTKSK 360
 Db 372 BCAMTEILMVGIMWSVEHPGLSFAPNLLDRNQRCAVEGLVEIDMTVTTRRPMK 431
 QY 361 LEKEFEVTLKIALANSDSM-----HIEVEAVOKLQDYVLEALQDYERGOHN---E 409
 Db 432 LRGEFFLCKSIILINSVYTFLSSTLESLBEDTDLIHILDKIPTLVHFMAKSGSLQ 491
 QY 410 DPRRAGKIMTLPLIROTSTAVOHGFYNIKLEGQPMHLFLEMLEA 456
 Db 492 QORBLAQDILLISHIRHMSNKGMEHLYSMKCNAVYFLYDILLIEMLEDA 538

Search completed: August 3, 2004, 12:40:55
Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:38:06 ; Search time 41 Seconds
(without alignments)
3524.569 Million cell updates/sec

Title: US-10-054-841-4
Perfect score: 2388
Sequence: 1 MDSEVCLDPEPSFLHYEEL.....KLEGVPMKFLFLEMLKAV 458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_rnuc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2282	95.6	466	11	Q8CHC9 mus musculus
2	1756	73.5	434	11	Q8OVSI mus musculus
3	1751	73.3	433	11	Q8CTA6 mus musculus
4	1744	73.0	433	11	Q8CCV5 mus musculus
5	1106	46.3	262	4	Q9HCB2
6	1054.5	44.2	323	4	Q96102
7	936.5	39.2	299	4	Q8NAS8
8	739.5	31.0	484	5	Q9VSE9
9	739.5	31.0	484	5	Q8WS79
10	726	30.4	496	5	Q8WMB0
11	724	30.3	589	13	Q8AYH0
12	715	29.9	581	13	Q8UTM9
13	697.5	29.2	152	13	Q8GCB0
14	695	29.1	574	13	Q72KX8
15	691	28.9	129	4	Q9UNK4
16	690	28.9	554	13	Q90ZM8

ALIGNMENTS

17	689	28.9	620	13	Q90WH6	Q90wh6 clarias gar
18	684	28.6	583	13	Q80406	Q804g6 acanthopagr
19	683.5	28.6	542	13	Q90ZB6	Q90z66 squallus aca
20	682.5	28.6	431	6	Q95L13	Q95l13 ovis aries
21	681	28.5	564	13	Q90WV1	Q90wv1 carassius a
22	678	28.4	620	13	Q7SZ10	Q7sz10 fundulus he
23	677	28.4	578	13	Q8QHL0	Q8qhl0 paraliichthy
24	675.5	28.3	570	13	Q8QHL0	Q800q2 zoarces viv
25	664.5	27.8	553	13	Q7J2U2	Q72j23 brachydantio
26	663.5	27.8	553	13	Q8SKV7	Q8skv7 brachydantio
27	663.5	27.8	553	13	Q90WS8	Q90ws8 brachydantio
28	659.5	27.6	559	13	Q8JVB9	Q8jvb9 cyprinus ca
29	653.5	27.4	553	13	Q8AV62	Q8av62 brachydantio
30	648.5	27.2	671	13	Q7T2K7	Q72k77 halichoeres
31	646.5	27.1	458	13	Q9DDJ3	Q9ddj3 acanthopagr
32	645	27.0	553	13	Q804Q7	Q804q7 acanthopagr
33	637	26.7	562	13	Q8UW75	Q8uw75 oryzias lat
34	636	26.6	567	11	Q8B6G5	Q8b6g5 mus musculu
35	632	26.5	503	11	Q91Z86	Q91z86 mus musculu
36	629	26.3	592	13	Q90WS9	Q90ws9 mus musculu
37	624.5	26.2	565	13	Q80HK9	Q80hk9 paraliichthy
38	621	26.0	592	13	Q7T3U4	Q7t3u4 varicorhinu
39	621	26.0	601	13	Q98SM8	Q98sm8 brachydantio
40	610.5	25.6	436	5	Q8T3U5	Q8t3u5 candidia ba
41	600.5	25.1	436	5	Q8T3U5	Q8t3u5 candidia ba
42	595	24.9	522	5	Q8XX78	Q8xx78 brachistost
43	588.5	24.6	363	5	Q9UAF1	Q9uaf1 polyanthroc
44	586.5	24.6	499	6	Q9SMF0	Q9smf0 macaca arc
45	584	24.5	412	13	Q91840	Q91840 xenopus lae

RESULT 1

Q8CHC9	PRELIMINARY;	PRT;	466 AA.
AC	Q8CHC9;		
DT	01-MAR-2003 (TREMBLrel. 23, Created)		
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	MKIAA0832 protein (Fragment).		
GN	ESRRG OR MKIAA0832.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_Taxid=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RA	Okazaki N., Kikuno R., Ohara R., Inamoto S., Hara Y., Nagase T.,		
RA	Ohara O., Koga H.;		
RT	"Prediction of the coding sequences of mouse homologues of KIAA gene:		
RT	I. The complete nucleotide sequences of 100 mouse KIAA-homologous		
RT	cDNAs identified by screening of terminal sequences of cDNA clones		
RT	randomly sampled from size-fractionated libraries."		
RL	Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.		
MD	MDL: MGI:1347056; ESRG.		
DR	GO: GO:0005634; C:nucleus; IEA.		
DR	GO: GO:0003707; F:steroid hormone receptor activity; IEA.		
DR	GO: GO:0003700; F:transcription factor activity; IEA.		
DR	GO: GO:006355; F:regulation of transcription; DNA-dependent; IEA.		
DR	InterPro: IPR000536; Hormone_rec_1lg.		
DR	InterPro: IPR001723; Strdimn_receptor.		
DR	InterPro: IPR008946; Str_ncl_receptor.		
DR	InterPro: IPR001628; Znf_Csteroi.		
DR	Pfam: PF00104; hormone_rec; 1.		
DR	Pfam: PF00105; zf-C4; 1.		
DR	PRINTS: PR00398; STRDHOMONER.		
DR	PRINTS: PR00047; STRODPTNGER.		
DR	Prodom: PD000035; Znf_Csteroi. 1.		
DR	SMART: SM00430; HOL1; 1.		

DR SMART; SMO0399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 FT NON TER 1
 DR SEQUENCE 466 AA; 52006 MW; 1EB8C98B180D6FEF CRC64;
 SQ
 Query Match 95.6%; Score 2282; DB 11; Length 466;
 Best Local Similarity 99.8%; Pred. No. 1e-181;
 Matches 438; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DR 20 LCRSNKNDRIHDSGCSFIFKTEPSSPASLTDVNVHSPGSSDASGSYSTMNGHONGL 79
 DB 28 LCRSNKNDRIHDSGCSFIFKTEPSSPASLTDVNVHSPGSSDASGSYSTMNGHONGL 87
 QY 80 DSPPIYPSAPILIGSGPVAKLYDDCSSTIVEDPOTKCEYMLNSMPKRLCLVCGDIASGYH 139
 DB 88 DSPPIYPSAPILIGSGPVAKLYDDCSSTIVEDPOTKCEYMLNSMPKRLCLVCGDIASGYH 147
 QY 140 YGVASGEACKAFKRTIOGNIYSCPATNECEITRRRRSCQACRPMKCLKYGMLKEGYR 199
 DB 148 YGVASGEACKAFKRTIOGNIYSCPATNECEITRRRRSCQACRPMKCLKYGMLKEGYR 207
 QY 200 LDRVRGGRQKYRRIDAENSPYINPOLVOPAKKPYNKIVSHLLVADPEKTYAMPDPPTVPD 259
 DB 208 LDRVRGGRQKYRRIDAENSPYINPOLVOPAKKPYNKIVSHLLVADPEKTYAMPDPPTVPD 267
 QY 260 SDIKALTTICDLADRELVIYIGMAKHIPGFSTLSLADQMSLLOSAMWEILLIGVYRSLS 319
 DB 268 SDIKALTTICDLADRELVIYIGMAKHIPGFSTLSLADQMSLLOSAMWEILLIGVYRSLS 327
 QY 320 FEELVYADYIMDEQSKLAGLIDNNAILQIVKKYSMKLEKEFEVTLKALALANSDS 379
 DB 328 FEELVYADYIMDEQSKLAGLIDNNAILQIVKKYSMKLEKEFEVTLKALALANSDS 387
 QY 380 MHIEDVAVOKLQDVYHEALQDYEAQGHMEDPPRAGMLMTPLRLQTSKAVOHFYNIK 439
 DB 388 MHIEDVAVOKLQDVYHEALQDYEAQGHMEDPPRAGMLMTPLRLQTSKAVOHFYNIK 447
 QY 440 LEGKVPNHKLFLEMLEAKV 458
 DB 448 LEGKVPNHKLFLEMLEAKV 466

RESULT 2
 Q80VS1 PRELIMINARY; PRT; 434 AA.
 ID 080VS1
 AC 080VS1;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Similar to estrogen related receptor, beta (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC044858; AAH44858.1; --
 DR PIR; P07076; P07076.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001723; Stdrhm_receptor.
 DR InterPro; IPR008946; Str_ncl_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PRO0398; STRDHOMONER.
 DR PRINTS; PRO0447; STROIDFLNGER.

DR Prodom; PD000035; Znf_C4steroid; 1.
 DR SMART; SMO0430; HOLI; 1.
 DR SMART; SMO0399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 FT NON TER 1
 DR SEQUENCE 434 AA; 48484 MW; 3679E7337B5762DC CRC64;
 SQ
 Query Match 73.5%; Score 1756; DB 11; Length 434;
 Best Local Similarity 75.0%; Pred. No. 7e-138;
 Matches 327; Conservative 55; Mismatches 52; Indels 2; Gaps 1;

DR 23 RMSNDRRIHDSGCSFIFKTEPSSPASLTDVNVHSPGSSDASGSYSTMNGHONGL 82
 DB 1 RMSNDRRIHDSGCSFIFKTEPSSPASLTDVNVHSPGSSDASGSYSTMNGHONGL 60
 QY 83 PIYPSAPILIGSGPVAKLYDDCSSTIVEDPOTKCEYMLNSMPKRLCLVCGDIASGYH 142
 DB 61 PIYPSAPILIGSGPVAKLYDDCSSTIVEDPOTKCEYMLNSMPKRLCLVCGDIASGYH 118
 QY 143 ASCEACKAFKRTIOGNIYSCPATNECEITRRRRSCQACRPMKCLKYGMLKEGYR 202
 DB 119 ASCEACKAFKRTIOGNIYSCPATNECEITRRRRSCQACRPMKCLKYGMLKEGYR 178
 QY 203 VEGGRQKYRRIDAENSPYINPOLVOPAKKPYNKIVSHLLVADPEKTYAMPDPPTVPD 262
 DB 179 VEGGRQKYRRIDAENSPYINPOLVOPAKKPYNKIVSHLLVADPEKTYAMPDPPTVPD 238
 QY 263 KALTTICDLADRELVIYIGMAKHIPGFSTLSLADQMSLLOSAMWEILLIGVYRSLS 322
 DB 239 KALTTICDLADRELVIYIGMAKHIPGFSTLSLADQMSLLOSAMWEILLIGVYRSLS 298
 QY 323 ELVYADYIMDEQSKLAGLIDNNAILQIVKKYSMKLEKEFEVTLKALALANSDS 382
 DB 299 KLAIVEDYIMDEHSHRVLGLDLYRAILQVRRYKLYKEKEFEVTLKALALANSDS 358
 QY 383 EDVEAVOKLQDVYHEALQDYEAQGHMEDPPRAGMLMTPLRLQTSKAVOHFYNIK 442
 DB 359 EDVEAVOKLQDVYHEALQDYEAQGHMEDPPRAGMLMTPLRLQTSKAVOHFYNIK 418
 QY 443 KVPNHKLFLEMLEAKV 458
 DB 419 KVPNHKLFLEMLEAKV 434

RESULT 3
 Q8C7A6 PRELIMINARY; PRT; 433 AA.
 ID 08C7A6
 AC 08C7A6;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Estrogen related receptor.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Heart;
 RC MEDLINE=22354683; PubMed=12466651;
 RA The FANTOM Consortium;
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RT Nature 420:563-573 (2002).
 RL EMBL; AK052256; BA034898.1; --
 DR EMBL; AK052256; BA034898.1; --
 DR PIR; P07076; P07076.
 DR MGD; MGI:1346832; Estrrb.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

```

OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Rowen L., Madan A., Qin S., Baradaran L., Birditt B., Bloom S.,
 RA Dora M., Dickhoff R., Fleetwood P., Harrison G., Kaur A., Madan A.,
 RA Nesbitt R., Traicoff R., Hood L.;
 RA "Sequencing of human chromosome 14q24.3 region."
 RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 DR EMBL; AC016543; AAC29619.1; -.
 DR HSSP; P03372; IHQ.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR001723; Strnrm_receptor.
 DR InterPro; IPR008946; Str_ncl_receptor.
 DR InterPro; IPR01628; Znf_ncl_Csteroi.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR PRODOM; PD000035; Znf_C4steroi.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
 KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 FT NON_TER
 SQ SEQUENCE 262 AA; 28203 MW; 94FA8979AC2D43D3 CRC64;
 Query Match 46.3%; Score 1106; DB 4; Length 262;
 Best Local Similarity 78.0%; Pred. No. 5,76-84;
 Matches 206; Conservative 26; Mismatches 30; Indels 2; Gaps 2;
 QY 24 MSNKRHIDSGSFFITPESSPASLDSVNHSPGSSSDASGSYSSTMNHQGLDPSR 83
 DB 1 MSSDRLHIGSSCGSFITPESSPIDALSHSPSGSDASGFGALGTHANGGLSP 60
 QY 84 LYPANPILGSGVPKLYVDDCSSTVEDPOTKCEYMLNMPKRLCLVCGDIASGVHYVA 143
 DB 61 MFAAG-LGGT-PCKSYEDCAISGIMEDSAIKCEYMLAIRKLCVCGDIASGVHYVA 118
 QY 144 SCEACKAFKFTIOGNIYSCPATNECEITRRRKSQACRFMKCLKVGMKEGVRLDRV 203
 DB 119 SCEACKAFKFTIOGNIYSCPATNECEITRRRKSQACRFMKCLKVGMKEGVRLDRV 178
 QY 204 RGGQKRYRRIDANSPYLPOLVOPAKPKYKNTYSHLYVPEKTYAMPDPTVDSIK 263
 DB 179 RGGQKRYRRIDANSPYLPOLVOPAKPKYKNTYSHLYVPEKTYAMPDPTVDSIK 238
 QY 264 ALTLCDLADRELVIITGMAKHIP 287
 DB 239 ALTLCDLADRELVIITGMAKHIP 262
 RESULT 6
 Q96102 PRELIMINARY; PRT; 323 AA.
 ID Q96102
 AC Q96102
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 DR EMBL; BC007915; AA07915.1; -.

DR HSSP; P19793; ILBD.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR InterPro; IPR000536; Hormone_rec_119.
 DR InterPro; IPR001723; Strnrm_receptor.
 DR InterPro; IPR008946; Str_ncl_receptor.
 DR InterPro; IPR01628; Znf_C4steroi.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRODOM; PD000035; Znf_C4steroi.
 DR SMART; SM00430; HOL1; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00402; BPD_TRANSP INN MEMBER; 1.
 KW Hypothetical protein; DNA-binding; Metal-binding; Nuclear protein;
 KW Receptor; Transcription; Transcription regulation; Zinc; Zinc-finger.
 FT NON_TER
 SQ SEQUENCE 323 AA; 35341 MW; 75D5BA85958025DC CRC64;
 Query Match 44.2%; Score 1054.5; DB 4; Length 323;
 Best Local Similarity 63.7%; Pred. No. 1,56-79;
 Matches 202; Conservative 50; Mismatches 52; Indels 13; Gaps 2;
 QY 153 KRTIQNIYSCPATNECEITRRRKSQACRFMKCLKVGMKEGVRLDRVGGQKRYR 212
 DB 4 KRTIQSIYSCPATNECEITRRRKSQACRFMKCLKVGMKEGVRLDRVGGQKRYR 63
 QY 213 RLDNENPYLPOLVOP-----AKRPYKLYSHLYVPEKTYAMPDPTVDSIK 263
 DB 64 RPEVDPPFPFPFAGPLVAVAGGRKTAAPNALVSHLYVPEKTYAMPDPTVDSIK 123
 QY 264 ALTLCDLADRELVIITGMAKHIPGFTSLADQMSILQSAAMETLLIGVYRSLSFDE 323
 DB 124 AVATLCLDFDEIVVITSMKSTIFGSSLSQMSVQSVWMEVLYGVAGRSPLQDE 183
 QY 324 LVYADVIMEDQSKLAGLIDINNAIQLVKKYSMKLEKEEFTYKAIALANSQSHTE 383
 DB 184 LAFADVLIDEGARAAGLIGALLQVRLQLRLREERYVYLKALALANSQSHTE 243
 QY 384 DVEAVQKIQVILHRLQVYEAQ----HMDPRRAGKMLMTPLILRQSTKAVQHFYNIK 439
 DB 244 DAEAVEQLREALHEALILEYEAAGPGGGAERRRARRLLTILPLRQTAQKVLAFHYVK 303
 QY 440 LEGKVPMEKLFLEMLEA 456
 DB 304 LEGKVPMEKLFLEMLEA 320
 RESULT 7
 Q9N4S8 PRELIMINARY; PRT; 299 AA.
 ID Q9N4S8
 AC Q9N4S8
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to estrogen-related receptor alpha (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=cervix;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC

```

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR EMBL: BC033701; AAH33701.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006350; P:regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR000536; Hormone_rec_119.
DR InterPro: IPR001723; Steroid_receptor.
DR InterPro: IPR008946; Str_ncl_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHORMNER.
DR SMART: SM00430; HOL1; 1.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER_1.
KM DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
FT NON_TER
SQ SEQUENCE 299 AA; 32629 MW; 186AC8A78B149AD9 CRC64;

Query Match 39.2%; Score 936.5; DB 4; Length 299;
Best Local Similarity 61.4%; Pred. No. 9,5e-70;
Matches 181; Conservative 48; Mismatches 53; Indels 13; Gaps 2;

QY 175 RRRKSCQACRFMKCLKVGMLKEGVRLDVRGGRQYKRIIDANSPYLFNLQVOP----- 229
2 RRPKACQACRFCTCLVGMGLKEGVRLDVRGGRQYKRIIDANSPYLFNLQVOP----- 61

QY 230 ---AKPKYKIVSHLVAEPKIVAMPDPVPDSIKALTTLCDLADRELVYIIGAKK 285
62 GPRKTAAPVAVLVAEPKIVAMPDPVPDSIKALTTLCDLADRELVYIIGAKK 121

QY 286 IPEFSTLADQMSLLOSAMELILIGVYRSLSFDELVADYIMDEQSLAGLIDL 345
122 IPEFSSLSISDQMSVLOSAMELILIGVYRSLSFDELVADYIMDEQSLAGLIDL 181

QY 346 NNAIILDVKKYKMKKEKEFEVTLKAIALANSMSMHIEDVEAYOKLDVYHEALDYDEAG 405
182 GAALLDQVRLQALRFREYVILKALANSMSMHIEDVEAYOKLDVYHEALDYDEAG 241

QY 406 Q---HMEPRRAQKMLTLPRLRQSTKVAQHRYNKLKGVYMHKLFLEMEA 456
242 RAGPGGGAERRRAGRLTLTFLRLQTAQKVLAFHYGVKLEGVYPMKLFLEMEA 296

RESULT 8
QVSES9 PRELIMINARY; PRT; 484 AA.
AC QVSES9;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DE 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CG7404 protein (GH28308P) (Estrogen-related receptor).
GN BRR OR CG7404.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCB1_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

```

```

RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Aoril J.F., Agbayan A., An H.-T., Andrews-Plamko C., Baldwin D.,
RA Baillet R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.W., Cawley S., Dahle K., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Goddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Jiang X., Iain X.,
RA Iain X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris U., Moshirei A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massaman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lehman J., Wherry S., Talmadge C., Davis J.R., Kiss I., Eudy J.D.,
RA Chase B.A., Sunegi J.;
RT "Identification and characterization of the Drosophila ortholog of
RL vertebrate estrogen-related receptor gene and its product."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
DR EMBL: AE003556; AAF50473.1; -.
DR EMBL: AY051632; AAK91056.1; -.
DR EMBL: AF359420; AAL37553.1; -.
DR HSSP: P03372; IHCO.
DR Flybase: FBgn0035849; ERR.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006350; P:regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR InterPro: IPR000536; Hormone_rec_119.
DR InterPro: IPR001723; Steroid_receptor.
DR InterPro: IPR008946; Str_ncl_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHORMNER.
DR PRINTS: PR00047; STROIDFINGER.
DR PRODOM: PD000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; ZNF_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.

```

KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger;
 SQ SEQUENCE 484 AA; 53245 MW; 252E3AD07C84A31 CAC64;
 Query Match 31.0%; Score 739.5; DB 5; Length 484;
 Best Local Similarity 37.4%; Pred. No. 5.2e-53;
 Matches 189; Conservative 70; Mismatches 151; Indels 95; Gaps 13;

14 LHYEEL-----LCMSKDKRHIDSSCSFKTER-----SSPSLIDSVNHHSPG 59
 8 LHIQEVPTSPASCFSPSSKSTATQSGTNGKSSPSVSPERQLCSSTLSLDCDHNVLS 67
 60 GSSDA-SSSYSTWM-CHQNGDSEPLTPAPILGSGFVRKLYDCCSTIYEDPQKCE 117
 68 NDGSIKSSGTSGGGGGGGCGTSGNATNASAGSGSVR---DEL----- 111
 118 YLMSMPKRLCLVCDIAGSYHYVASCEBCKAFKRTIQGNIEXSCPATNECEITKRR 177
 112 -----RLCLVCGDVASGFHYGVASCEBCKAFKRTIQGNIEXYCEANNECEITKRR 164
 178 KSCQACRPMKCLAKVMKEGVRLDVRGGRQYKRIANSPYLNQVPAKPEY--- 234
 165 KACQACRQCKLMLMKEGVRLDVRGGRQYKRR--NPVNSYQTMQLYQSNTSLCD 222
 235 NKIVSHLVABEPEKIYAM-PPTVPDSDI----- 262
 223 VKILEVNSYEPDALSVOPTPPQVHTTSTINDENASSSGSITLSSSVTPNGTCIFONNN 282
 263 -----KALTTLCDLADRELVITIAKHIIPGSTSLADQSLLOSANMEITLGVYRS 317
 283 NNDEPHEILSVLSDIYDKELVSVIGWAKOIPGFDLPINDQMLLOVSAEITLQITERS 342
 318 LSFEBELVADYDYMDEQSKLAGLLDNNALLOVKYKSKMKEKEPVTLKATIALNS 377
 343 LPENGKLCFADVWMDHILAKCGYTEFYHCVQIAQNERISPREEYTLKALLANLC 402
 378 DSMHIEDVEAVOKLDVLEALQDY-----EACQHMEDPRRACKMLMTLPLRLQSTIK 430
 403 DIL-LDDQSSLRAPFDITLINSINDVYLLRHSSANVSHQ-----QLLLPLSLNQADI 455
 431 AVQHFNKIKLEGVPMKLEMLE 455
 456 LRFPRGJARDEVITMKKLEMLE 480

RESULT 9
 Q8WS79 PRELIMINARY; PRT; 496 AA.
 ID Q8WS79;
 AC Q8WS79;
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Estrogen-related receptor splice variant (CG7404-PB).
 GN ERR OR CG7404.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lehman J., Cherry S., Talmadge C., Davis J.R., Rudy J.D., Chase B.A.,
 RA Kiss I., Smegi J.;
 RA "Identification and characterization of the Drosophila ortholog of
 RA vertebrate estrogen-related receptor gene and its product";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBD databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abiril J.F., Aghayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chertis J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.B., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Itali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Li X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spiter B., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodgett W., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Celinker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Barzon J., An H., Baldwin D., Barzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Drenek D., Farfan D.,
 RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Itali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svitskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBD databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Mista S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Rungall C.J., Lewis S.E.;
 RA "Annotation of Drosophila melanogaster genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBD databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBD databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBD databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 DR EMBL, AF359421; AAL37554.1; -.

DR EMBL; AB003556; AAM12015.1; -
 DR FlyBase; FBgn0035849; ERR.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR000536; Hormone_rec_1lg.
 DR InterPro; IPR001723; Steroid_receptor.
 DR InterPro; IPR008946; Str_ncl_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STROHOMNER.
 DR PRINTS; PR00047; STRODINGER.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOL1; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 SQ SEQUENCE 496 AA; 54529 MW; 5D67E5A8B9DCB32 CRC64;

Query Match 31.0%; Score 739.5; DB 5; Length 496;
 Best Local Similarity 37.4%; Pred. No. 5,4e-53;
 Matches 189; Conservative 70; Mismatches 151; Indels 95; Gaps 13;

14 LHYEEL-----LCMSNKRHIDSSCSFKTEP-----SSPALSIVNHSFG 59
 20 LHKQEVTPASASCPSPSKSTATOSGNGKLSBSVSPEROLCSSTLSLSDHNVLS 79
 60 GSSPA-GSSYSTNM-GHONGIDSPPIYSPAPILIGSSGPRKLYDDCSSTIVEDPQKCE 117
 80 NDSGLKSGSGSGNGGGGGGGTSGGNATNAGAGSSSVR---DEL----- 123
 118 YMLSMKRLCLVCGDIASGYHGVASCEACKAFKRTIGNIEYSCPATNECETRRR 177
 124 -----RRLVCGDVASGFHYGVASCEACKAFKRTIGNIEYTCANNECINNRR 176
 178 KSCQACEFMKCLKVGMKEGRLDRVGRGROKRRIDAENSPIYLNQVOPAKKY--- 234
 177 KACQACFPQCKLMGMKEGRLDRVGRGROKRR---NPVSNSTQTMQLYOSNTSLCD 234
 235 NKIVSHLVAPEKTIAM-DEPTVPDSI----- 262
 235 VKLLEVINSYEPALSVQTPPVHTTSTINDASSSSGSKLESSVYTPNGCTIQNN 294
 263 -----KALTLCDLARELVIIIGNAKHIPGSTLSLADQMSILQSAAMEILLGVYRS 317
 295 NNDPNEILVSLDIYKEIVSVIGMAKQIPGFIDLPIDQMKLLQVSWAEILLQLTPRS 354
 318 LSEFEDELVAADYIMDEQSKLAGLDLNNAILQVKKYKSMLEKEEFTTKALALANS 377
 355 LFNNGKLCFATVMDMDHLAKEGYEFYHCQVQIAKORNERISPREEEYLLKALLANC 414
 378 DSMHIEDVAOVOKLQDVHEALQDY-----EAGQMEDPRRAGKMLTLPIYRQTSRK 430
 415 DTL-LDDQSLEAFRDTIINSINDVYLLRHSSAVSHQ-----QLLLSLRQADI 467
 431 AVQHFNITLSEKVPMMKLFEMLE 455
 466 LRRFMRGIARDVITMKLLEMLE 492

RESULT 10
 Q8UWB0
 AC Q8UWB0 PRELIMINARY; PRT; 587 AA.
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Estrogen receptor.
 GN ER.

OS Caiman crocodilus (Spectacled caiman) (Caiman sclerops).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauia; Crocodylidae; Alligatorinae; Caiman.
 OX NCBI_TaxID=8499;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21490797; PubMed=11604222;
 RA Sumida K., Ooe N., Saito K., Kaneko H.;
 RT "Molecular cloning and characterization of reptilian estrogen receptor
 cDNAs";
 RL Mol. Cell. Endocrinol. 183:33-39(2001).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 DR EMBL; AB055220; BAB79436.1; -
 DR GO; GO:0016585; C:chromatin remodeling complex; ISS.
 DR GO; GO:0016020; C:membrane; ISS.
 DR GO; GO:0030284; F:estrogen receptor activity; ISS.
 DR GO; GO:0030235; F:nitric-oxide synthase regulator activity; ISS.
 DR GO; GO:0003520; P:estrogen receptor signaling pathway; ISS.
 DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; ISS.
 DR InterPro; IPR000536; Hormone_rec_1lg.
 DR InterPro; IPR001723; Steroid_receptor.
 DR InterPro; IPR008946; Str_ncl_receptor.
 DR InterPro; IPR001628; Znf_C4steroid.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF02159; Oest_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STROHOMNER.
 DR PRINTS; PR00047; STRODINGER.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOL1; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 SQ SEQUENCE 587 AA; 66514 MW; DE172B837FEF513 CRC64;

Query Match 30.4%; Score 726; DB 13; Length 587;
 Best Local Similarity 38.2%; Pred. No. 9,2e-52;
 Matches 175; Conservative 80; Mismatches 143; Indels 60; Gaps 14;

38 FIKTEPSSPALSLDVSVMHSPGG-----SSPASSTYTMNGHONGIDSPPIYSPAPILGG 93
 104 FLQTAP-----QLSPFHHHSQVPPYLINDQSP-----GMBPAASTFYRSADSRHQ 153
 94 SGVRLKYDDCSSTIVEDPQTKCYMINSMPK-RLCLVCGDIASGYHGVASCEACKAF 152
 154 SGRER-----MSST-----SEKASLMESTKETRYCAVNCNDVSGHYGVASCEGCKAF 203
 153 KRTIGNIEYSCPATNECETITRRRSQAQCPMKCLKVGMKEGVRLDRVGRGRO-KYK 211
 204 KSIQGHNDYPCGATNQCTIDKNRRRSQAQCLRKCYEVMGKGIKOR-RGRRLKQK 262
 212 RR-----IDANSP-----YLNQVOPAKK-----PYNKIVSHLVAPEKTY 250
 263 RQREODANRGTAAMETPTLMTSPVITKTKKNSPALSTIYAQMSALAEAPPIY 322
 251 AMPDPVPSDIKALFTTCDLADRELVIIGNAKHIPGSTLSLADQMSILQSAAMEILL 310
 323 SEYDNRPRPNEASMMTLTLNLADRELVIINWAKVPRVVDLTLDQVHLEFCAMIBILM 382
 311 LGVYRSLSFEDELVAADYIMDEQSK-LAGLDLNNAILQVKKYKSMLEKEEFTTKAL 369
 383 IGLVRSKHEHFKCLFAFNLLIDRVQGCVGMEVIFPMLLATAARFMMNLQGEFVCL 442
 370 KALALAN-----SDSMHIEDVAOVOKLQDVHEALQDYEAQGM---EDPRAGKML 418
 443 KSIILNSGVYTPSLSTLSKLSBKDYHRYVDKTIYLIHMASSGLSQOQHRLLAQIL 502
 419 MTLPLRLQTSKAVQHFYNYTLSEKVPMMKLFEMLEA 456
 503 LILSHIRMSNKGMEHLNMYCKKNVPLYLILLEMIDA 540

QY 200 LDEVRGRO-KYKRRD-----AENSP-YINPOLVPAKK-----PYNKIV 238
 Db 248 KOR-RGRMLKXKRGDELDGNNAYVYEAHTITLMPSPMKIKSKNSPALSTLAEQWY 306
 QY 239 SHLLVAEPKITYAMPDPYVSDIKALITLDCDLADRELVIYIGAKHI PGFSTLADQM 298
 Db 307 SALIDAEPIVYSEYDPSFSEASVMTLITLADRELVIYIMAKVPGFVLDALADQV 366
 QY 299 SLIQSAMWELLIGVYRSLSFEDEIVYADVIMEDOSK-LAGLIDNNALITLVKXK 357
 Db 367 HLEFCAMWELLIGVYRSLSFEDEIVYADVIMEDOSK-LAGLIDNNALITLVKXK 357
 QY 358 SMLEKEEFVTLKALAN-----SDSMIEDVBAVQKLDVYHEALQDYAGQM- 408
 Db 427 MNNOGEEFVTLKALAN-----SDSMIEDVBAVQKLDVYHEALQDYAGQM- 408
 QY 409 --BDPRAGKMTLPLRLQTSKAVQHPYNTKLGKPMHFLFLEMEA 456
 Db 487 LQOQHRLAQILIMLSHIRMKNKMEHLYNKKCNVPLVLDLLEMDA 536

RESULT 13

Q8G80 PRELIMINARY; PRT; 152 AA.
 ID 08G80; O8G80;
 AC 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DB Bacteroides receptor-related receptor (fragment).
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 CX NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Westerlund P., Olsson P.-E.;
 RT "Cloning of zebrafish nuclear hormone receptors";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 DR EMBL: AF495875; C:nucleus; IEA.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO: GO:0006350; P:transcription; IEA.
 DR InterPro: IPR001723; Stdrhm_receptor.
 DR InterPro: IPR008946; Str_ncl_receptor.
 DR InterPro: IPR001628; Znf_ncl_receptor.
 DR Pfam: PF00105; Zf-C4; 1.
 DR PRINTS: PR00398; STRDHOMONER.
 DR PRODOM: PD000035; Znf_C4steroid; 1.
 DR SMART: SM00359; Znf_C4; 1.
 KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KM Transcription regulation; Zinc; Zinc-finger.
 FT NON TER 1
 FT TER 152
 SQ SEQUENCE 152 AA; 17189 MW; A97FALBDPE091B1 CRC64;

Query Match 29.2%; Score 697.5; DB 13; Length 152;
 Best Local Similarity 87.4%; Pred. No. 3.1e-50;
 Matches 132; Conservative 5; Mismatches 7; Indels 7; Gaps 1;
 QY 146 EACCAFRTIQTGNIYSCPAINECEITTKRRKSCQACRFMKCLKVGMKESGVRLDRVYG 205
 Db 2 KGCAFFERTIQNIYSCPAINECEITTKRRKSCQACRFMKCLKVGMKESGVRLDRVYG 61
 QY 206 GRCYKRIAPENSPYINPOLVPAKKY-----NKIVSHLLVAEPKITYAMPDPYV 258
 Db 62 GRCYKRIAPENSPYINPOLVPAKKY-----NKIVSHLLVAEPKITYAMPDPYV 121
 QY 259 DSDIKALITLDCDLADRELVIYIGAKHIPGF 289

Db 122 DSDIKALITLDCDLADRELVIYIGAKHIPGF 152

RESULT 14

Q7T2K8 PRELIMINARY; PRT; 574 AA.
 ID 07T2K8; Q7T2K8;
 AC 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DB Estrogen receptor alpha.
 OS Haliccoeres tenuispinus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Labroidae;
 OC Labridae; Halichoeres.
 CX NCBI_TaxId=217853;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSE-Ovary,
 RA Park J.-G., Kim S.-U.;
 RT "Molecular cloning and expression of two estrogen receptor subtypes in
 RT proteogynous wrasse, Halichoeres tenuispinus";
 RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY305026; AAP72178.1; -.
 KW Receptor.
 SQ SEQUENCE 574 AA; 63661 MW; 34626475977971B1 CRC64;

Query Match 29.1%; Score 695; DB 13; Length 574;
 Best Local Similarity 33.9%; Pred. No. 3.4e-49;
 Matches 168; Conservative 90; Mismatches 169; Indels 68; Gaps 12;

QY 9 PEFSLIAYEELLGMSNKHIDSCSFPITKEPPSPASLTDVYVNHSPGSSDASGSV 68
 Db 31 PTLVSLSTQGYSAALDTHGQPSDSISQ--LGSGPSPLVFVPSPLSP-----F 80
 QY 69 SSTNMGCHNGDSPILYSAFILLGSGPVKLY---DQCSSTVED---PQTCEVYVNS 122
 Db 81 MHLFPHHILYESTFTVYVSSVSSQSGISREHCGTSDBSYSWEGSAGAAACCFMAKE 140
 QY 123 MPRKLCVQDGIASGYHYGVASCEACAFKRTIQTGNIYSCPAINECEITTKRRKSCQ 182
 Db 141 M-RYCVNSDYVASGYHYGVASCEACAFKRTIQTGNIYSCPAINECEITTKRRKSCQ 198
 QY 183 CFPMKCLKVGMKESGVRLDRV-----GGRQYKRRY 214
 Db 199 CRLRCYEVGMKGGVRRDGRVLRDRTGTSDKNGSKDRQRTVPRQGRKSSSV 258
 QY 215 DAEVNSPYINPOLVPAKKPYNKIVSHLLVAEPKITYAMPDPYVSDIKALITLDCDLAD 274
 Db 259 GGGKSPVIS-----MPPQVLLIQAGSPILICSRQKLSRPYEVMTMTLTSMTDR 310
 QY 275 ELVYIIGAKHIPGSTSTLADQMSILOSAMWELLIGVYRSLSFEDEIVYADYIMDE 334
 Db 311 ELVYIIGAKHIPGSTSTLADQMSILOSAMWELLIGVYRSLSFEDEIVYADYIMDE 370
 QY 335 DQSK-LAGLIDNNALITLVKXKSMLEKEEFVTLKALANSDSMH-----IDV 385
 Db 371 SEDGCVGMAEFIDMLATTSFRMLKLPSEFVCLAIILNNGAFSFCGTGMEPHDN 430
 QY 386 EAVQKLDVYHEALQDY--EAG--QHMEPDRAGKMTLPLRLQTSKAVQHPYNTKLG 441
 Db 431 EAVQKLDVYHEALQDY--EAG--QHMEPDRAGKMTLPLRLQTSKAVQHPYNTKLG 489
 QY 442 GRYVPMKCLFLEMEA 456
 Db 490 NKVPLVLDLLEMDA 504
 RESULT 15
 Q9UNU4 PRELIMINARY; PRT; 129 AA.
 ID Q9UNU4

AC 09JUN04: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Estrogen receptor related protein 3 (Fragment).
GN ERR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9357798; PubMed=10428842;
RT Hong H., Yang L., Stallcup M.R.;
RT "Hormone-independent transcriptional activation and coactivator
binding by novel orphan nuclear receptor ERR3."
RT J. Biol. Chem. 274:22618-22626 (1999).
DR EMBL; AF117255; AAD48370.1; -
DR GO: GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 129 129
SQ SEQUENCE 129 AA; 13990 MW; 7BE487F716CCD48 CRC64;
Query Match 28.9%; Score 691; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 8.6e-50;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSVELCLPESFSLHYERELLCRMNSKDRHIDSCSSFFIKTEPSSPASLTDVNVHSPGG 60
DB 1 MDSVELCLPESFSLHYERELLCRMNSKDRHIDSCSSFFIKTEPSSPASLTDVNVHSPGG 60
QY 61 SSDAGSYSTWNGHONGIDSPILYPSAPIIGSGPVRLYDCCSTIYEDPQTKCEYML 120
DB 61 SSDAGSYSTWNGHONGIDSPILYPSAPIIGSGPVRLYDCCSTIYEDPQTKCEYML 120
QY 121 NSMPKRLCL 129
DB 121 NSMPKRLCL 129

Search completed: August 3, 2004, 12:41:49
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:39:31 ; Search time 19 Seconds

(without alignments)
1244.458 Million cell updates/sec

Title: US-10-054-841-4

Perfect score: 2388

Sequence: 1 MDSVELCLPESFSLHYEBEL.....KLEGVPMHKLFTLEMLKAV 458

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5E.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2388	100.0	458	3	US-09-141-000-4
2	2265	94.8	435	3	US-09-040-508-2
3	2265	94.8	435	4	US-09-500-654-2
4	2139	89.6	418	3	US-09-141-000-6
5	1799	75.3	500	3	US-09-141-000-2
6	1775	74.3	431	2	US-08-836-620A-19
7	1212.5	50.8	518	2	US-08-836-620A-18
8	716.5	30.0	596	2	US-08-836-620A-16
9	704	29.5	595	3	US-08-764-870-12
10	704	29.5	595	3	US-08-980-115-12
11	702	29.4	591	2	US-08-836-620A-17
12	702	29.4	595	2	US-08-836-620A-15
13	702	29.4	595	2	US-09-041-886-35
14	685	28.7	595	4	US-08-453-998-2
15	661	27.7	410	6	5223606-5
16	661	27.7	484	2	US-08-836-620A-14
17	655	27.4	548	2	US-08-836-620A-5
18	655	27.4	548	3	US-09-139-617-1
19	655	27.4	548	4	US-09-561-741A-1
20	651	27.3	484	4	US-09-558-795-1
21	651	27.3	484	2	US-08-836-620A-13
22	648	27.0	485	2	US-08-836-620A-2
23	644.5	27.0	477	4	US-09-608-088-25
24	644.5	27.0	495	2	US-09-608-088-5
25	597	25.0	411	3	US-08-836-620A-3
26	597	25.0	411	4	US-09-040-508-4
27	596.5	25.0	467	1	US-09-500-654-4
					Sequence 4, Appl1

28	596.5	25.0	467	4	US-08-216-592A-6	Sequence 6, Appl1
29	596.5	25.0	467	5	PCT-US91-00393-4	Sequence 4, Appl1
30	589	24.7	462	1	US-08-336-408B-2	Sequence 2, Appl1
31	589	24.7	462	3	US-08-764-870-6	Sequence 6, Appl1
32	589	24.7	462	3	US-08-980-115-6	Sequence 6, Appl1
33	589	24.7	462	5	PCT-US91-00393-2	Sequence 2, Appl1
34	577.5	24.2	463	4	US-08-216-592A-8	Sequence 8, Appl1
35	574	24.0	533	4	US-08-216-592A-4	Sequence 4, Appl1
36	573.5	24.0	384	2	US-08-836-620A-15	Sequence 15, Appl1
37	573.5	24.0	416	4	US-09-608-088-6	Sequence 6, Appl1
38	573.5	24.0	418	4	US-09-608-088-21	Sequence 21, Appl1
39	573	24.0	533	1	US-07-952-800-2	Sequence 2, Appl1
40	572.5	24.0	463	1	US-08-336-408B-6	Sequence 6, Appl1
41	572.5	24.0	463	5	PCT-US91-00393-6	Sequence 6, Appl1
42	571	23.9	525	3	US-08-764-870-7	Sequence 7, Appl1
43	571	23.9	525	3	US-08-980-115-7	Sequence 7, Appl1
44	560	23.5	446	1	US-07-952-800-4	Sequence 4, Appl1
45	558	23.4	448	4	US-08-216-592A-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1									
US-09-141-000-4									
; Sequence 4, Application US/09141000									
; Patent No. 6054295									
; GENERAL INFORMATION:									
; APPLICANT: Chen, Pang									
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR									
; FILE REFERENCE: 19999Y									
; CURRENT APPLICATION NUMBER: US/09/141,000									
; NUMBER OF SEQ ID NOS: 30									
; SOFTWARE: FASTSEQ for Windows Version 3.0									
; SEQ ID NO 4									
; LENGTH: 458									
; TYPE: PRT									
; ORGANISM: Human									
US-09-141-000-4									
Query Match									
Best Local Similarity 100.0%; Score 2388; DB 3; Length 458;									
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MDSVELCLPESFSLHYEBEL	CMNKNDRHIDSSCSFTKTESSPASLDSVNHSPGG	60					
DB	1	MDSVELCLPESFSLHYEBEL	CMNKNDRHIDSSCSFTKTESSPASLDSVNHSPGG	60					
QY	61	SSDASGSYSTWNGHNGLDSPPLYPSPAPITLGGSGVRKLYDDCSSTIVEDPOTKCEYML	120						
DB	61	SSDASGSYSTWNGHNGLDSPPLYPSPAPITLGGSGVRKLYDDCSSTIVEDPOTKCEYML	120						
QY	121	NSMPKRLCLVCGDIAAGHYGVASCACACAFRRRTIQGNIEYSCPATNECETIKRRKSC	180						
DB	121	NSMPKRLCLVCGDIAAGHYGVASCACACAFRRRTIQGNIEYSCPATNECETIKRRKSC	180						
QY	121	NSMPKRLCLVCGDIAAGHYGVASCACACAFRRRTIQGNIEYSCPATNECETIKRRKSC	180						
DB	121	NSMPKRLCLVCGDIAAGHYGVASCACACAFRRRTIQGNIEYSCPATNECETIKRRKSC	180						
QY	181	QACRFPKCLKVGMLKXGVRIDRVRGROKRYRIIDANSPYANPOLVOPAKKRYNKIVSH	240						
DB	181	QACRFPKCLKVGMLKXGVRIDRVRGROKRYRIIDANSPYANPOLVOPAKKRYNKIVSH	240						
QY	181	QACRFPKCLKVGMLKXGVRIDRVRGROKRYRIIDANSPYANPOLVOPAKKRYNKIVSH	240						
DB	181	QACRFPKCLKVGMLKXGVRIDRVRGROKRYRIIDANSPYANPOLVOPAKKRYNKIVSH	240						
QY	241	LIVAEPEKTIYAMDPTVPDSIDIKALITLDCDLADRELVIIGMAKHIPGFTLSIADQMSL	300						
DB	241	LIVAEPEKTIYAMDPTVPDSIDIKALITLDCDLADRELVIIGMAKHIPGFTLSIADQMSL	300						
QY	241	LIVAEPEKTIYAMDPTVPDSIDIKALITLDCDLADRELVIIGMAKHIPGFTLSIADQMSL	300						
DB	241	LIVAEPEKTIYAMDPTVPDSIDIKALITLDCDLADRELVIIGMAKHIPGFTLSIADQMSL	300						
QY	301	LQSAWKEIILIGVYSLSFEDELVYADYIMEDOSKLAGLIDINNAIIQVKKYKSMK	360						
DB	301	LQSAWKEIILIGVYSLSFEDELVYADYIMEDOSKLAGLIDINNAIIQVKKYKSMK	360						
QY	301	LQSAWKEIILIGVYSLSFEDELVYADYIMEDOSKLAGLIDINNAIIQVKKYKSMK	360						
DB	301	LQSAWKEIILIGVYSLSFEDELVYADYIMEDOSKLAGLIDINNAIIQVKKYKSMK	360						
QY	361	LEKEEPEVTIKALIALNSDSMHIEDVAVOGLQVLEALQDYVAGOMEDPRRAGKGLMT	420						
DB	361	LEKEEPEVTIKALIALNSDSMHIEDVAVOGLQVLEALQDYVAGOMEDPRRAGKGLMT	420						
QY	361	LEKEEPEVTIKALIALNSDSMHIEDVAVOGLQVLEALQDYVAGOMEDPRRAGKGLMT	420						
DB	361	LEKEEPEVTIKALIALNSDSMHIEDVAVOGLQVLEALQDYVAGOMEDPRRAGKGLMT	420						

QY 421 LPLRQSTKAVOHFYNIKLKGVPMHKLFLMLEAKV 458
DB 421 LPLRQSTKAVOHFYNIKLKGVPMHKLFLMLEAKV 458

RESULT 2

US-09-040-508-2
; Sequence 2, Application US/09040508
; Patent No. 6069239
; GENERAL INFORMATION:
; APPLICANT: Mathias, Stephen L.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,508
; FILING DATE: 17-MAR-1998
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9705451.4
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: GB 9803289.9
; FILING DATE: 16-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: CH-30008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-040-508-2

Query Match 94.8%; Score 2265; DB 3; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.7e-225;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MSNKHIDSSCSFPTKTPSSPASLTDVNHSPGSSDASGSYSTNMGHONGIDSP 83
DB 1 MSNKHIDSSCSFPTKTPSSPASLTDVNHSPGSSDASGSYSTNMGHONGIDSP 60
QY 84 LYPSPAILGGSGPVKRLYDDCSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 143
DB 61 LYPSPAILGGSGPVKRLYDDCSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 120
QY 144 SCEBACAFPKRTIQGNIIEYSCPATNECITTKRRRSCQACRFMKCLKVGMKEGVYLDV 203
DB 121 SCEBACAFPKRTIQGNIIEYSCPATNECITTKRRRSCQACRFMKCLKVGMKEGVYLDV 180
QY 204 RGRGRQYKRIIDAENSPYINPOLVOPAKKPYNKIVSHLVAEPEKIYAMPDPTVPDSIK 263
DB 181 RGRGRQYKRIIDAENSPYINPOLVOPAKKPYNKIVSHLVAEPEKIYAMPDPTVPDSIK 240
QY 264 ALTTICDLADRELIVYIGMAKHIPGFSTLSLADQMSILOSAMWELLIGVYRSLSFEDE 323

DB 241 ALTTICDLADRELIVYIGMAKHIPGFSTLSLADQMSILOSAMWELLIGVYRSLSFEDE 300
QY 324 LYYADYIMDEDSQSLAGLIDNNAILQVKKYKSMKEKEFFVTLKAIALANSMSHIE 383
DB 301 LYYADYIMDEDSQSLAGLIDNNAILQVKKYKSMKEKEFFVTLKAIALANSMSHIE 360
QY 384 DVEAVQKLDVLAHEALQDYEAQGHMEDPRRAGKMLMTPLRLQSTKAVOHFYNIKLKGV 443
DB 361 DVEAVQKLDVLAHEALQDYEAQGHMEDPRRAGKMLMTPLRLQSTKAVOHFYNIKLKGV 420
QY 444 VPMHKLFLMLEAKV 458
DB 421 VPMHKLFLMLEAKV 435

RESULT 3

US-09-500-654-2
; Sequence 2, Application US/09500654
; Patent No. 6359116
; GENERAL INFORMATION:
; APPLICANT: Stephen L. Mathias
; TITLE OF INVENTION: No. 6359116e1 Compounds
; FILE REFERENCE: GP-30008-D1
; CURRENT APPLICATION NUMBER: US/09/500,654
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 9705451.4
; EARLIER FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: 9803289.9
; EARLIER FILING DATE: 1998-02-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-500-654-2

Query Match 94.8%; Score 2265; DB 4; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.7e-225;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MSNKHIDSSCSFPTKTPSSPASLTDVNHSPGSSDASGSYSTNMGHONGIDSP 83
DB 1 MSNKHIDSSCSFPTKTPSSPASLTDVNHSPGSSDASGSYSTNMGHONGIDSP 60
QY 84 LYPSPAILGGSGPVKRLYDDCSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 143
DB 61 LYPSPAILGGSGPVKRLYDDCSTIVEDPQTKCEYMLNSMPKRLCLVCGDIASGYHYGA 120
QY 144 SCEBACAFPKRTIQGNIIEYSCPATNECITTKRRRSCQACRFMKCLKVGMKEGVYLDV 203
DB 121 SCEBACAFPKRTIQGNIIEYSCPATNECITTKRRRSCQACRFMKCLKVGMKEGVYLDV 180
QY 204 RGRGRQYKRIIDAENSPYINPOLVOPAKKPYNKIVSHLVAEPEKIYAMPDPTVPDSIK 263
DB 181 RGRGRQYKRIIDAENSPYINPOLVOPAKKPYNKIVSHLVAEPEKIYAMPDPTVPDSIK 240
QY 264 ALTTICDLADRELIVYIGMAKHIPGFSTLSLADQMSILOSAMWELLIGVYRSLSFEDE 323
DB 241 ALTTICDLADRELIVYIGMAKHIPGFSTLSLADQMSILOSAMWELLIGVYRSLSFEDE 300
QY 324 LYYADYIMDEDSQSLAGLIDNNAILQVKKYKSMKEKEFFVTLKAIALANSMSHIE 383
DB 301 LYYADYIMDEDSQSLAGLIDNNAILQVKKYKSMKEKEFFVTLKAIALANSMSHIE 360
QY 384 DVEAVQKLDVLAHEALQDYEAQGHMEDPRRAGKMLMTPLRLQSTKAVOHFYNIKLKGV 443
DB 361 DVEAVQKLDVLAHEALQDYEAQGHMEDPRRAGKMLMTPLRLQSTKAVOHFYNIKLKGV 420
QY 444 VPMHKLFLMLEAKV 458
DB 421 VPMHKLFLMLEAKV 435

RESULT 4
US-09-141-000-6
; Sequence 6, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; TITLE OF INVENTION: RECEPTOR PROTEINS
; FILE REFERENCE: 19999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-6

Query Match 89.6%; Score 2139; DB 3; Length 418;
Best Local Similarity 100.0%; Pred. No. 1,8e-212;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSVELCLPESFSLHYEHELICRMSNKKDRHIDSSCSFTTEPSSPASLTDVYNHSPGG 60
DB 1 MDSVELCLPESFSLHYEHELICRMSNKKDRHIDSSCSFTTEPSSPASLTDVYNHSPGG 60
QY 61 SSDASGSSSTNMGNHONGLDSPPLYPSPAPILGSGGVRKLYDDCSSTIYEDPOTKCEYML 120
DB 61 SSDASGSSSTNMGNHONGLDSPPLYPSPAPILGSGGVRKLYDDCSSTIYEDPOTKCEYML 120
QY 121 NSMPKRLCLVCGDIAGGYHGVASCAACAFKFKRTIQGNIEYSCPATNCEITKRRKSC 180
DB 121 NSMPKRLCLVCGDIAGGYHGVASCAACAFKFKRTIQGNIEYSCPATNCEITKRRKSC 180
QY 181 QACRFKCLKVGMLKEGVRLDRVRGROKRYKRIIDANSPYINPOLVOPAKKRYNKTIVSH 240
DB 181 QACRFKCLKVGMLKEGVRLDRVRGROKRYKRIIDANSPYINPOLVOPAKKRYNKTIVSH 240
QY 241 LVAABPEKLYAMPDPTVPDSDIKALTTLCIADRELVIITGAKHIFGFTSLADQMSI 300
DB 241 LVAABPEKLYAMPDPTVPDSDIKALTTLCIADRELVIITGAKHIFGFTSLADQMSI 300
QY 301 LOSAMWEIILIGVYRSLSEDELVTAADYIMDEDSKLAGLIDNNAILQIVKXYKSKM 360
DB 301 LOSAMWEIILIGVYRSLSEDELVTAADYIMDEDSKLAGLIDNNAILQIVKXYKSKM 360
QY 361 LKKEEFVTLKATALANDSMHIEDVEAYOKLODVLEALQDYFAGQHME 409
DB 361 LKKEEFVTLKATALANDSMHIEDVEAYOKLODVLEALQDYFAGQHME 409

RESULT 5
US-09-141-000-2
; Sequence 2, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Fang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; TITLE OF INVENTION: RECEPTOR PROTEINS
; FILE REFERENCE: 19999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-2

Query Match 75.3%; Score 1799; DB 3; Length 500;
Best Local Similarity 78.1%; Pred. No. 3.6e-177;
Matches 339; Conservative 50; Mismatches 43; Indels 2; Gaps 2;

QY 24 MSNKKDRHIDSSCSFTTEPSSPASLTDVYNHSPGGSSPASGSSYSTNMGNHONGLDSP 83
DB 1 MSSDDRLHSGSCSFTTEPSSPSGIDALSHSPGSSDASGAGFALDTNANGLDSP 60
QY 84 LYPAPILGSGGVRKLYDDCSSTIYEDPOTKCEYMLNSMPKRLCLVCGDIAGGYHGA 143
DB 61 MFAGAG-LGGT-PKRSYEDCAAGIMEDSAIKETYMALPKRLCLVCGDIAGGYHGA 118
QY 144 SCACAFKFKRTIQGNIEYSCPATNCEITKRRKSCQACRFKCLKVGMLKEGVRLDRV 203
DB 119 SCACAFKFKRTIQGNIEYSCPATNCEITKRRKSCQACRFKCLKVGMLKEGVRLDRV 178
QY 204 RGRORRYKRIIDANSPYINPOLVOPAKKRYNKTIVSHLVAABPEKLYAMPDPTVPDSDIK 263
DB 179 RGRORRYKRIIDANSPYINPOLVOPAKKRYNKTIVSHLVAABPEKLYAMPDPTVPDSDIK 238
QY 264 ALTTLCIADRELVIITGAKHIFGFTSLADQMSILOSAMWEIILIGVYRSLSEDE 323
DB 239 ALTTLCIADRELVIITGAKHIFGFTSLADQMSILOSAMWEIILIGVYRSLSEDE 298
QY 324 LVAADYIMDEDSKLAGLIDNNAILQIVKXYKSKMLKEKEEFVTLKATALANDSMHIE 383
DB 299 LVAADYIMDEDSKLAGLIDNNAILQIVKXYKSKMLKEKEEFVTLKATALANDSMHIE 358
QY 384 DVEAYOKLODVLEALQDYFAGQHMEPPRAGKMLTLPLRQTSKAVHFNKLEBK 443
DB 359 DVEAYOKLODVLEALQDYFAGQHMEPPRAGKMLTLPLRQTSKAVHFNKLEBK 418
QY 444 VPMHKLFLMLELAK 457
DB 419 VPMHKLFLMLELAK 432

RESULT 6
US-08-836-620A-19
; Sequence 19, Application US/0883620A
; Patent No. 5958710
; GENERAL INFORMATION:
; APPLICANT: Orphan receptor
; TITLE OF INVENTION: Orphan receptor
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,620A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/03933
; FILING DATE:
; APPLICATION NUMBER: GB 9518272.1
; FILING DATE: 08-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9605550.4
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9607532.0
; FILING DATE: 11-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9609576.5
; FILING DATE: 08-MAY-1996
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; ORIGINAL SOURCE:

ORGANISM: Homo sapiens
US-08-836-620A-19

Query Match 74.3%; Score 1775; DB 2; Length 431;
Best Local Similarity 76.4%; Pred. No. 8.6e-175;
Matches 331; Conservative 51; Mismatches 49; Indels 2; Gaps 1;

```

QY 24 MSNDRHIDSSCSFITEPSSPASLNDVNVHSPGSSDASGSSYSSMMNHQGLDSSP 83
DB 1 MSSDRHLGSSCSGFITEPSSPASLNDVNVHSPGSSDASGSSDASGALGTHANGDSSP 60
QY 84 LYPAPILGSGSPRKLYDDCSSTIVEDPQYKCEYMLNSMPEKRLCLVCGDIASGVHGV 143
DB 61 MFAAGL--GGNPKRSYEDCTSGIMEDSAIKCEYMLNAIKRLCLVCGDIASGVHGV 118
QY 144 SCEACKAFKRTIOGNIEYSCPAINECHITRRRKSQACPFMKCLKYGMKEGYRLDRV 203
DB 119 SCEACKAFKRTIOGNIEYSCPAINECHITRRRKSQACPFMKCLKYGMKEGYRLDRV 178
QY 204 RGGQKXRRIDANSPLYNPOLVQAPAKPYNKIVSHLVAEPEKIVAMPDPTVPDSDIK 263
DB 179 RGGQKXRRIDANSPLYNPOLVQAPAKPYNKIVSHLVAEPEKIVAMPDPTVPDSDIK 238
QY 264 ALTLICDLADELVIIGMAKHIPGFTLSLADQMSLLQSAAMEILLIGVYRSLSPFDE 323
DB 239 ALTLICDLADELVIIGMAKHIPGFTLSLADQMSLLQSAAMEILLIGVYRSLSPFDE 298
QY 324 LVADVYIMDEDSKLAGLIDLNNAILOVKKYKSMKEKEFPVTLKAIALANSDSMAIE 383
DB 299 LAVEDYIMDEHSEVLGLLELYRALIQLVRRYKDLKVEKEFWLKAIALANSDSMYIE 358
QY 384 DVEAVOKLQDYLHEALQDYEAQGHEDPRRAGKMLMTPLRLRQSTKVAOHFYNYKLEGX 443
DB 359 NLEAVOKLQDYLHEALQDYELQSHREBPRRAGKMLMTPLRLRQSTKVAOHFYNYKLOK 418
QY 444 VPMKLFLEMLEA 456
DB 419 VPMKLFLEMLEA 431

```

RESULT 7

US-08-836-620A-18
Sequence 18, Application US/08836620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT: Orphan receptor
TITLE OF INVENTION: Orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,620A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/03933
FILING DATE:
APPLICATION NUMBER: GB 9518272.1
FILING DATE: 08-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 960550.4
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9607532.0
FILING DATE: 11-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9609576.5
FILING DATE: 08-MAY-1996
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids

TYPE: amino acid
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-836-620A-18

Query Match 50.8%; Score 1212.5; DB 2; Length 518;
Best Local Similarity 53.5%; Pred. No. 1.9e-116;
Matches 250; Conservative 67; Mismatches 85; Indels 65; Gaps 9;

```

QY 22 CRMSNKRHIDSSCS-----FITEPSSPASLNDVNVHSPGSSD-----A 64
DB 85 CPTLPEPQVYLSANSSQVVGIEPLYIKAEPPSP-----DSPKSSETEPEPVALA 135
QY 65 SG-SYSTMGHOGGLDSSPPLYPGAPILGSGSPRKLYDDCSSTIVEDPQYKCEYMLNSM 123
DB 136 PGAPPTKLPQHKKEEDGE-----GAP-----GEQGGKLVLSL 171
QY 124 PKRLCLVCGDIASGVHGVASCEACKAFKRTIOGNIEYSCPAINECHITRRRKSQAC 183
DB 172 PKRLCLVCGDVASGVHGVASCEACKAFKRTIOGSIIEYSCPASNECHITRRRKSQAC 231
QY 184 RPKMCLKYGMKEGYRLDRVNGRQKYYRRIDAENSPYLNQVLP-----AKPY 234
DB 232 RFTKIRYGMKEGYRLDRVNGRQKYYRRIDAEVDPPEPGFPAGPLAVAGPRTAAV 291
QY 235 NKIVSHLVAEPEKIVAMPDPTVPDSDIKALTLICDLADELVIIGMAKHIPGFTLSL 294
DB 292 NALVSHLVAEPEKIVAMPDPTVPDSDIKALTLICDLADELVIIGMAKHIPGFTLSL 351
QY 295 ADQMSLLQSAAMEILLIGVYRSLSPFDELYADVYIMDEDSKLAGLIDLNNAILOVYK 354
DB 352 SDQMSVLSQVMEVYLVGVAQRSLPLQDELAFAEEDVLIBEGARAAGIGELGALLQVLR 411
QY 355 KYKSMKEKEFPVTLKAIALANSDSMAIEDEYEA-QLQDYLHEALQDYEAQGHEDPR 409
DB 412 RLQARLREERYVLKAIALANSDSVHIEDEPRMSSCEKLEHLLEBEGARAGPGGGA 471
QY 410 DPRRAGKMLMTPLRLRQSTKVAOHFYNYKLEKGYPMKLFLEMLEA 456
DB 472 ERRRAGRLTLPLRLRQSTKVAOHFYNYKLEKGYPMKLFLEMLEA 518

```

RESULT 8

US-08-836-620A-16
Sequence 16, Application US/08836620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT: Orphan receptor
TITLE OF INVENTION: Orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,620A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/03933
FILING DATE:
APPLICATION NUMBER: GB 9518272.1
FILING DATE: 08-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 960550.4
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9607532.0
FILING DATE: 11-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9609576.5
FILING DATE: 08-MAY-1996

; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 596 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; ORIGINAL SOURCE:
 ; ORGANISM: Rattus rattus
 ; US-08-836-620A-16

Query Match 30.0%; Score 716.5; DB 2; Length 596;
 Best Local Similarity 39.2%; Pred. No. 5,1e-65;
 Matches 168; Conservative 82; Mismatches 128; Indels 51; Gaps 13;

```

QY 73 NGHQNLDSPPY-----PSAPLIGSGSP---VEKLYDDCSSTIVEPQKCE-----YMLN 121
DB 129 HGHQ-----VPYLENEPSAYAVRDGPAPFRNSDNRRQNGRERLSSSSKGNMIMES 183
QY 122 SMPKRLCLVCGDIASGYHYGVAACECAKAFKRTIQGNIYSCPATNECEITKRRKSCQ 181
DB 184 AERTYCAVCNDYASGYHYGVAWSCEGCAFFKRSIQGHNDYCPATNOCTIDKRRKSCQ 243
QY 182 ACRPKCKLVGMLKEGVRLDRVGRGQ-KYKR-RIDAE-----SPYLNPLQ 226
DB 244 ACRLRKCYEVMKMGKIRKDR-RGGRMLKHKRQDDLEGRNEMGTSGDMRAANLWPSPLY 302
QY 227 VQPAK-----PYNKIVSHLVAEPKRYAMPDPVDSIDKALITLCLADRELVI 279
DB 303 IKHTKNSPALSLTADQVNSALIDAPPLIYSEYDPSRPSASMMGLITNLADRELVM 362
QY 280 IGMATHIGFSTLSADQMSLISAMMEITLIGVYRSLSPEDLVAVDYIMDDQSK- 338
DB 363 INMARVGFQDNLHDQVHLBECAMLEILMIGVWRSMERPKLFFAPNILLDRNQKC 422
QY 339 LAGLLDNNALIQLVKRYKSKLEKEEFVTLKATLAN-----SSMHTEDVEAVOK 390
DB 423 VEGWVEIFDMLATSSRRFMNLOGEFVCLKSIILNSGYTFLSTLSLEEKDHR 482
QY 391 LQVLAHEALQDY--EAGQHMEDP--RRAGKMLMTPLPLRQSTKAVOHEVYNIKLKGVMH 447
DB 483 VLDKINDTLILHMAKAGLTLOQQRRLAQLLILSHIRHMSNKGMEHLVYKCKXNVVFLY 542
QY 448 KLFLEMEA 456
DB 543 DLTLEMDA 551

```

RESULT 9
 ; US-08-764-870-12
 ; Sequence 12, Application US/08764870
 ; Patent No. 6236946
 ; GENERAL INFORMATION:
 ; APPLICANT: Scanlan, Thomas S
 ; APPLICANT: Baxter, John D
 ; APPLICANT: Fletterick, Robert J
 ; APPLICANT: Wagner, Richard L
 ; APPLICANT: Kushner, Peter J
 ; APPLICANT: Apriletti, James W
 ; APPLICANT: West, Brian
 ; TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
 ; TITLE OF INVENTION: Binding Domains
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Cooley Godward
 ; STREET: Five Palo Alto Square, 3000 El Camino Real
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/764,870
 ; FILING DATE: 13-DEC-1996
 ; CLASSIFICATION: 530
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/008,540
 ; FILING DATE: 13-DEC-1995
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/008,543
 ; FILING DATE: 13-DEC-1995
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/008,606
 ; FILING DATE: 14-DEC-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Nakamura, Jackie N
 ; REGISTRATION NUMBER: 35,966
 ; REFERENCE/DOCKET NUMBER: UCAL-246/01US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650)843-5000
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 595 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-764-870-12

Query Match 29.5%; Score 704; DB 3; Length 595;
 Best Local Similarity 41.2%; Pred. No. 1e-63;
 Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

```

QY 119 MANSMPKRLCLVCGDIASGYHYGVAACECAKAFKRTIQGNIYSCPATNECEITKRRK 178
DB 176 MESAKERTYCAVCNDYASGYHYGVAWSCEGCAFFKRSIQGHNDYCPATNOCTIDKRRK 235
QY 179 SCQACRFMKCKLVGMLKEGVRLDRVGRGQ-KYKR-RIDAE-----SPYLN 223
DB 236 SCQACRLRKCYEVMKMGKIRKDR-RGGRMLKHKRQDDLEGRNEMGTSGDMRAANLWPS 294
QY 224 POLVQPAK-----PYNKIVSHLVAEPKRYAMPDPVDSIDKALITLCLADREL 276
DB 295 PLMTKRKNSLALSTADQVNSALIDAPPLIYSEYDTPRPSBASMMGLITNLADREL 354
QY 277 VLIIGMAKHIPGSTLSADQMSLISAMMEITLIGVYRSLSPEDLVAVDYIMDDQ 336
DB 355 VHMIMAKRYVGFQDNLHDQVHLBECAMLEILMIGVWRSMERPKLFFAPNILLDRNQ 414
QY 337 SK-LAGLLDNNALIQLVKRYKSKLEKEEFVTLKATLAN-----DSMH 381
DB 415 GKCEWVEIFDMLATSSRRFMNLOGEFVCLKSIILNSGYTFLSTLSLEEKD 474
QY 382 IEDVEAVOKLQVLAHEALQDYENGQHMEDP--RRAGKMLMTPLPLRQSTKAVOHEVYNIKL 440
DB 475 IHRV--LDKITDILHMA--KAGLTLOQQRRLAQLLILSHIRHMSNKGMEHLVYKCK 530
QY 441 EGVYPMHKLPLEVLEMA 456
DB 531 KNVPELYDLLEMDA 546

```

RESULT 10
 ; US-08-980-115-12
 ; Sequence 12, Application US/08980115
 ; Patent No. 6266622
 ; GENERAL INFORMATION:
 ; APPLICANT: Scanlan, Thomas S.
 ; APPLICANT: Baxter, John D.
 ; APPLICANT: Fletterick, Robert J.
 ; APPLICANT: Wagner, Richard L.
 ; APPLICANT: Kushner, Peter J.
 ; APPLICANT: Apriletti, James W.
 ; APPLICANT: West, Brian L.

APPLICANT: Shiao, Andrew K.
 TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
 FILE REFERENCE: UGAL-246/02US
 CURRENT APPLICATION NUMBER: US/08/980,115
 CURRENT FILING DATE: 1997-11-26
 EARLIER APPLICATION NUMBER: 08/764,870
 EARLIER FILING DATE: 1996-12-13
 EARLIER APPLICATION NUMBER: 60/008,606
 EARLIER FILING DATE: 1995-12-14
 EARLIER APPLICATION NUMBER: 60/008,543
 EARLIER FILING DATE: 1995-12-13
 EARLIER APPLICATION NUMBER: 60/008,540
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO: 12
 LENGTH: 595
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: DOMAIN
 LOCATION: (287)..(549)
 OTHER INFORMATION: minimal ligand binding domain
 US-08-980-115-12

Query Match 29.5%; Score 704; DB 3; Length 595;

Best Local Similarity 41.2%; Pred. No. 1e-63;

Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

QY 119 MNSMRKRLCTVCGDIASGHHGVASCEACAFKFTTIOGNIYSCPATNECITRRRK 178
 DB 176 MESAKETRYCAVNDYASGYHGVSCGCKAFKSIQGHNDYMCPTATNCTIDKRRK 235
 QY 179 SCQACFMKCLKYGMLKEGRLDRVGRQ-KYKR-RIDAEN-----SPYLN 223
 DB 236 SCQACLRKCYEYGMKGGIRKOR-RGRMLKHKRQDGEGRGEVGSAGDMRAANLWPS 294
 QY 224 POLVOPAKR-----PYNKIVSHLVABPEKTYAMPDPVPSPDIKALTTLODLADREL 276
 DB 295 PLMIKRSKNSLALSLADQVNSALDAEPILYSEYDPRPSEASMGILTNLADREL 354
 QY 277 VLIIGNAKHIPGFTSLADQMSLLOSAMWEIILGVYRSLFEDELVYADYIMDEQ 336
 DB 355 VHMIMNAKRVGFVDLTHQVHLBCAMWILMIGVWMSMEHPKILFAPNILLDRNQ 414
 QY 337 SK-LAGLIDNNAILOLVKXKSKLKEKEPVTIKATALANS-----DSMH 381
 DB 415 GKCEGVWEIIFDMLATSSRRMNLOGEFVCLKSIILNSGVYTELSSTLKSLEEKDH 474
 QY 382 IEDVEAVQKLODVILHEALQDYEAQGHMEDP-RRAGKMLMTLPLRQSTKAVOHFYNIKL 440
 DB 475 IHRV--LDKITDTLHMA--KAGLTLQOQHRLAQLLILSHIRHMSNKGMEHLYSKWC 530
 QY 441 EGKVPMEKLFLEMLEA 456
 DB 531 KNVVPYDILLLEMDA 546

RESULT 11

US-08-936-620A-17

Sequence 17, Application US/08836620A

Patent No. 5958710

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Orphan receptor

NUMBER OF SEQUENCES: 19

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/936,620A

FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP96/03933
 FILING DATE: 08-SEP-1995
 APPLICATION NUMBER: GB 9518272.1
 FILING DATE: 08-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9605550.4
 FILING DATE: 15-MAR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9607532.0
 FILING DATE: 11-APR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9609576.5
 FILING DATE: 08-MAY-1996
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 591 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-08-936-620A-17

Query Match 29.4%; Score 702; DB 2; Length 591;

Best Local Similarity 41.2%; Pred. No. 1.6e-63;

Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

QY 119 MNSMRKRLCTVCGDIASGHHGVASCEACAFKFTTIOGNIYSCPATNECITRRRK 178
 DB 176 MESAKETRYCAVNDYASGYHGVSCGCKAFKSIQGHNDYMCPTATNCTIDKRRK 235
 QY 179 SCQACFMKCLKYGMLKEGRLDRVGRQ-KYKR-RIDAEN-----SPYLN 223
 DB 236 SCQACLRKCYEYGMKGGIRKOR-RGRMLKHKRQDGEGRGEVGSAGDMRAANLWPS 294
 QY 224 POLVOPAKR-----PYNKIVSHLVABPEKTYAMPDPVPSPDIKALTTLODLADREL 276
 DB 295 PLMIKRSKNSLALSLADQVNSALDAEPILYSEYDPRPSEASMGILTNLADREL 354
 QY 277 VLIIGNAKHIPGFTSLADQMSLLOSAMWEIILGVYRSLFEDELVYADYIMDEQ 336
 DB 355 VHMIMNAKRVGFVDLTHQVHLBCAMWILMIGVWMSMEHPKILFAPNILLDRNQ 414
 QY 337 SK-LAGLIDNNAILOLVKXKSKLKEKEPVTIKATALANS-----DSMH 381
 DB 415 GKCEGVWEIIFDMLATSSRRMNLOGEFVCLKSIILNSGVYTELSSTLKSLEEKDH 474
 QY 382 IEDVEAVQKLODVILHEALQDYEAQGHMEDP-RRAGKMLMTLPLRQSTKAVOHFYNIKL 440
 DB 475 IHRV--LDKITDTLHMA--KAGLTLQOQHRLAQLLILSHIRHMSNKGMEHLYSKWC 530
 QY 441 EGKVPMEKLFLEMLEA 456
 DB 531 KNVVPYDILLLEMDA 546

RESULT 12

US-09-041-886-35

Sequence 35, Application US/09041886

Patent No. 6233872

GENERAL INFORMATION:

APPLICANT: Bredesen, Dale E.

TITLE OF INVENTION: Rabitazadeh, Sharroz

TITLE OF INVENTION: Proapoptotic Peptides, Dependence

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell & Flores LLP

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-886-35

Query Match 29.4%; Score 702; DB 3; Length 595;
Best Local Similarity 41.2%; Pred. No. 1.6e-63;
Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

119 M L N S M P R L C L V C D I A S G H Y G V A S C E A C A F P K R T I O G N I E Y S C P A T N E C E I T K R R K 178
176 M E S A K E T R Y C A V C N D Y A S G H Y G V A S C E G C A F P K S I Q H N D Y M C P A T N O C T I D K R R K 235
179 S C A C R P M K C L K Y G M L K E G V R L D R V R G R Q - K Y K R - R I D A E N - - - - - S P Y L N 223
236 S C A C R L R K C Y E V M M G G I R K D R - R G R M L K H R Q D D E G R G E V S A G D M R A A N I M P S 294
224 P O L V O P A K K - - - - - P Y N K I V S H L V A E P E K I Y A M P D P T V P S D I K A L T T L C D L A D R E L 276
295 P L M I K S K K S N S L A L S T A D Q M S A L D A B P P I L Y S E Y D P R P F S E A S M M G L T N I A D R E L 354
277 V L I I G A K H I P G F S T L S L A D Q M S I L O S A M M E I L L G V Y R S L S F E D E L V Y A D Y I M D E Q 336
355 V H I M N A K R V P G V D L T L H D V H L E C A M L E I M I G L V W R S M E H P V K L F A P N I L D R N Q 414
337 S K - L A G L D L N N A I L O L V K Y K S M K L E K E E F V T L K A I A L A N S - - - - - D S M H 381
415 G K C V E G M V E I F P M L A T S S R F R M N L Q G E E F V C I K S I I L N S G V Y T F L S T L K S L E K D H 474
382 I E D V E A V O K I O D V I H E A L O D Y E A G O H M E D P - R R A G K M L M T L P L R Q S T K A V O H F Y N I K L 440
475 I H R V - - L D K I T D T L I H L M A - - K A G L T L Q O O H R L A Q L L I L S H I R M S N K G M E H L Y S M K C 530
441 E G K V P M H K L F L E M L E A 456
531 K N V V P L Y D L L E M L D A 546

RESULT 13
US-08-453-998-2
Sequence 2, Application US/08453998
Patent No. 644438
GENERAL INFORMATION:
APPLICANT: CHAMSON, PIERRE
APPLICANT: METZGER, DANIEL
APPLICANT: WHITE, JOHN
TITLE OF INVENTION: METHOD FOR THE PREPARATION OF A PROTEIN
TITLE OF INVENTION: BY YEASTS USING AN INDICIBLE SYSTEM, VECTORS AND
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CUSHMAN, DARY & CUSHMAN
STREET: 1100 NEW YORK AVE., N.W.

CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,998
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/161,064
FILING DATE: 03-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: CHAPIN, MARILYN K.
REGISTRATION NUMBER: 35,843
REFERENCE/DOCKET NUMBER: 1037/98493
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3711
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-453-998-2

Query Match 29.4%; Score 702; DB 4; Length 595;
Best Local Similarity 41.2%; Pred. No. 1.6e-63;
Matches 155; Conservative 75; Mismatches 103; Indels 43; Gaps 10;

119 M L N S M P R L C L V C D I A S G H Y G V A S C E A C A F P K R T I O G N I E Y S C P A T N E C E I T K R R K 178
176 M E S A K E T R Y C A V C N D Y A S G H Y G V A S C E G C A F P K S I Q H N D Y M C P A T N O C T I D K R R K 235
179 S C A C R P M K C L K Y G M L K E G V R L D R V R G R Q - K Y K R - R I D A E N - - - - - S P Y L N 223
236 S C A C R L R K C Y E V M M G G I R K D R - R G R M L K H R Q D D E G R G E V S A G D M R A A N I M P S 294
224 P O L V O P A K K - - - - - P Y N K I V S H L V A E P E K I Y A M P D P T V P S D I K A L T T L C D L A D R E L 276
295 P L M I K S K K S N S L A L S T A D Q M S A L D A B P P I L Y S E Y D P R P F S E A S M M G L T N I A D R E L 354
277 V L I I G A K H I P G F S T L S L A D Q M S I L O S A M M E I L L G V Y R S L S F E D E L V Y A D Y I M D E Q 336
355 V H I M N A K R V P G V D L T L H D V H L E C A M L E I M I G L V W R S M E H P V K L F A P N I L D R N Q 414
337 S K - L A G L D L N N A I L O L V K Y K S M K L E K E E F V T L K A I A L A N S - - - - - D S M H 381
415 G K C V E G M V E I F P M L A T S S R F R M N L Q G E E F V C I K S I I L N S G V Y T F L S T L K S L E K D H 474
382 I E D V E A V O K I O D V I H E A L O D Y E A G O H M E D P - R R A G K M L M T L P L R Q S T K A V O H F Y N I K L 440
475 I H R V - - L D K I T D T L I H L M A - - K A G L T L Q O O H R L A Q L L I L S H I R M S N K G M E H L Y S M K C 530
441 E G K V P M H K L F L E M L E A 456
531 K N V V P L Y D L L E M L D A 546

RESULT 14
5223606-5
Patent No. 5223606
APPLICANT: BLAUDIN DE THE, HUGHES, MARCITO, AGNES, TIOLLAIS,
PIERRE, DEIRAN, ANNE
TITLE OF INVENTION: STEROID/THYROID HORMONE RECEPTOR-RELATED
PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCYLLULAR CARCINOMA

NUMBER OF SEQUENCES: 11
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/134,130
 FILING DATE: 17-DEC-1987
 PRIOR APPLICATION DATA:
 SEQ ID NO: 5
 LENGTH: 410

5223606-5

Query Match 28.7%; Score 685; DB 6; Length 410;
 Best Local Similarity 41.7%; Pred. No. 5e-62;
 Matches 153; Conservative 73; Mismatches 97; Indels 44; Gaps 11;

QY 128 CAVGDIASGVHYGASCEACKAFKRTIQANIYSCPAITBCEITTKRRRSQACRPMK 187
 DB 1 CAVCNDAVSGHYGVWSCGCKAFKRSIQGHNDYMCATNOCITDKRRRSQACRLRK 60
 QY 188 CLKVGMKEGVRLDVRGRGQ-KYKR-RIDEN-----SPYLPOLVQPAKK 232
 DB 61 CYEVGMKGGRKQK-RGRMKLKHQRDDGEGEVGSAGMRAANLWPSLMIKRSKX 119
 QY 223 -----PYNKIVSHLVAEPEKIYAMPDPVPSDIALATTLCDLADRELVIIIGAKH 285
 DB 120 NSLALSLADQVNSALDAEP-ILYSEYDPTRPSEASWMLLTMLADRELVHMINWAKR 178
 QY 286 IPGFSTLSIADQMSILQSAWMEILIGVYRSLSFEDELIVYADDYIMDEDSK-LAGLLD 344
 DB 179 VGFVDLTLDHQVHLLECAMELILMIGVWRSMEHVKLLFAPNLLIDRQKCVGEWVE 238
 QY 345 LNNALIQLVKKYKSMLEKEFEVTLKAILANS-----DSMHIEDVEAVOK 390
 DB 239 IPDMILATSSRRMMNLQGEFEVCLKSLILNSGVTTFTSLTKSLBEKDHIHRV-LDK 296
 QY 391 LGDVLHEALQDYEAGQMEDP-RRAGKMLMTPLRLQSTKAVQHFYNIKLKGKVPMMK 449
 DB 297 ITDTLHLMA--KAGLTLOQGHRLAQLLLISHIHMSKMGHEHLYSMCKKVVVLVYL 354
 QY 450 FLEMLEA 456
 DB 355 LLEMLDA 361

RESULT 15
 US-08-836-620A-14
 ; Sequence 14, Application US/08836620A
 ; Patent No. 5958710

GENERAL INFORMATION:

APPLICANT: TITLE OF INVENTION: Orphan receptor
 NUMBER OF SEQUENCES: 19
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/836,620A
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP96/03933
 FILING DATE:
 APPLICATION NUMBER: GB 9518272.1
 FILING DATE: 08-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9605550.4
 FILING DATE: 15-MAR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9607532.0
 FILING DATE: 11-APR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9609576.5
 FILING DATE: 08-MAY-1996
 INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
 LENGTH: 484 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: Mus musculus
 US-08-836-620A-14

Query Match 27.7%; Score 661; DB 2; Length 484;
 Best Local Similarity 38.6%; Pred. No. 2e-59;
 Matches 150; Conservative 79; Mismatches 118; Indels 42; Gaps 11;

QY 91 LGSGGPVARKLYDDCSSTIVEDPQTCCEYMLNSMKRKLAVGCDIASGVHYGASCEACKA 150
 DB 83 LGSGG-----CASP-VTSPSTKRD-----AHFCAVCSDIASGVHYGVWSCGCKA 126
 QY 151 FFKRTIQANIYSCPAITBCEITTKRRRSQACRPMKCLKVGMKEGVRLDVRGRGQKY 210
 DB 127 FFKRSIQGHNDYICPAITNOCITDKRRRSQACRLRKCYEVGMVCGSRRERC-GYRIVR 185
 QY 211 KRRLDAENSPYLN-----POLVQPAKPYNKIVSHLVAEPEKIYAMPDPVTP 258
 DB 186 RQRSASQVHCLNKAKRTSGHTPRVKEILLNSLP-EQLVLTLEAPPNV-LVSRPSWP 243
 QY 259 DSDIALATTLCDLADRELVIIIGAKHIGPSTLSIADQMSILQSAWMEILIGVYRSLS 318
 DB 244 FTEASMMMSLTGLADKEIVHMGAKKIPGEVELSLDQVRLBSCGMHEVIMVGLMRSI 303
 QY 319 SFEDELIVYADDYIMDEDSK-LAGLLDNNALIQLVKKYKSMLEKEFEVTLKAILANS 377
 DB 304 DHPGKLIFAPDLVLDREBKCVEGILEIFDMLATTARRELKLOHKEVLCVAMAILNS 363
 QY 378 DSMHI-----EDVEAVOKLQDVLH---EALQDYEAGQMEDP---RAGKMLMTPLRLQ 427
 DB 364 SMVHLTASQGEABSSKLTLLINAVTDALVWVYSKRISQOQSVRIANLMLLSHVRHI 423
 QY 428 STKAVQHFYNIKLKGKVPMMKLEMLEA 456
 DB 424 SNKGMEHLLSMCKKVVVPYDILLEMLNA 452

Search completed: August 3, 2004, 12:42:46
 Job time: 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:40:36 ; Search time 45 Seconds

(without alignments)
3192.595 Million cell updates/sec

Title: US-10-054-841-4

Perfect score: 2388
Sequence: 1 MDSVELCLPESFSLHYEEL.....KLEGVPMKFLTEMLKAV 458

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2388	100.0	458	US-10-054-841-4	Sequence 4, Appl1
2	2388	100.0	458	US-10-355-218-1	Sequence 1, Appl1
3	2139	89.6	418	US-10-054-841-6	Sequence 6, Appl1
4	1799	75.3	500	US-10-054-841-2	Sequence 2, Appl1
5	1775	74.3	431	US-10-278-481-19	Sequence 19, Appl1
6	1259	52.7	486	US-10-297-653-1	Sequence 1, Appl1
7	1258.5	52.7	519	US-10-211-239-1	Sequence 1, Appl1
8	1212.5	50.8	518	US-10-278-481-18	Sequence 18, Appl1
9	1073	44.9	515	US-10-093-463-202	Sequence 202, Appl1
10	947	39.7	230	US-10-355-218-8	Sequence 8, Appl1
11	755	31.6	234	US-10-355-218-9	Sequence 9, Appl1
12	726	30.4	587	US-10-450-484A-1	Sequence 1, Appl1
13	716.5	30.0	596	US-10-278-481-16	Sequence 16, Appl1
14	715	29.9	581	US-10-451-768A-1	Sequence 1, Appl1
15	713.5	29.9	595	US-10-052-092-31	Sequence 31, Appl1

16	713.5	29.9	595	14	US-10-437-107-31	Sequence 31, Appl1
17	711.5	29.8	595	14	US-10-052-092-30	Sequence 30, Appl1
18	711.5	29.8	595	14	US-10-437-107-30	Sequence 30, Appl1
19	707	29.6	595	15	US-10-148-835-8	Sequence 8, Appl1
20	704	29.5	595	9	US-09-853-033-2	Sequence 2, Appl1
21	704	29.5	595	15	US-10-148-835-1	Sequence 1, Appl1
22	704	29.5	595	15	US-10-148-835-3	Sequence 3, Appl1
23	704	29.5	595	15	US-10-148-835-9	Sequence 9, Appl1
24	703	29.4	595	15	US-10-148-835-5	Sequence 5, Appl1
25	702	29.4	591	14	US-10-278-481-17	Sequence 17, Appl1
26	702	29.4	595	9	US-09-933-267A-2	Sequence 2, Appl1
27	702	29.4	595	10	US-09-952-680A-10	Sequence 10, Appl1
28	702	29.4	595	13	US-10-096-710-1	Sequence 1, Appl1
29	702	29.4	595	13	US-10-081-563-2	Sequence 2, Appl1
30	702	29.4	595	14	US-10-052-092-9	Sequence 9, Appl1
31	702	29.4	595	14	US-10-052-092-13	Sequence 13, Appl1
32	702	29.4	595	14	US-10-052-092-14	Sequence 14, Appl1
33	702	29.4	595	14	US-10-207-655-61	Sequence 61, Appl1
34	702	29.4	595	14	US-10-177-293-128	Sequence 128, Appl1
35	702	29.4	595	14	US-10-157-899A-55	Sequence 55, Appl1
36	702	29.4	595	14	US-10-437-107-9	Sequence 9, Appl1
37	702	29.4	595	14	US-10-437-107-13	Sequence 13, Appl1
38	702	29.4	595	14	US-10-437-107-14	Sequence 14, Appl1
39	702	29.4	595	14	US-10-095-373A-2	Sequence 2, Appl1
40	702	29.4	595	14	US-10-392-274-2	Sequence 2, Appl1
41	702	29.4	595	15	US-10-148-835-6	Sequence 6, Appl1
42	702	29.4	595	15	US-10-144-198-42	Sequence 42, Appl1
43	701	29.4	595	15	US-10-148-835-2	Sequence 2, Appl1
44	701	29.4	595	15	US-10-148-835-7	Sequence 7, Appl1
45	697	29.2	595	15	US-10-148-835-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-10-054-841-4
Sequence 4, Application US/10054841
Publication No. US20030119100A1
GENERAL INFORMATION:
APPLICANT: Chen, Fang
TITLE OR INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
FILE REFERENCE: 19999YCA
CURRENT APPLICATION NUMBER: US/10/054, 841
CURRENT FILING DATE: 2002-01-23
PRIOR APPLICATION NUMBER: 09/487, 379
PRIOR FILING DATE: 2000-01-18
PRIOR APPLICATION NUMBER: 09/141, 000
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/078, 633
PRIOR FILING DATE: 1998-03-19
PRIOR APPLICATION NUMBER: 60/062, 902
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/057, 090
PRIOR FILING DATE: 1997-08-27
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 458
TYPE: PRT
ORGANISM: Human
US-10-054-841-4
Query Match 100.0%; Score 2388; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.4e-215;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSVELCLPESFSLHYEELCRMNKRHIDSSCSFFIKTEPSPASTLTVNVHSPGG 60
DB 1 MDSVELCLPESFSLHYEELCRMNKRHIDSSCSFFIKTEPSPASTLTVNVHSPGG 60
QY 61 SSDASGSYSTMNQNGDSSPLPYSPAPILGSGFVRKLYDDCSSTIVEDPQTKCEYML 120

```
Db 61 SSDASGYSSTWNGHNGDSDPLYPSPAPILGSSGVPKRLYDDCSSTIVEDPOTKCEYML 120
Qy 121 NSMPKRLCLVCGDIAGGYHGVASCEACAFPERKTIQGNIEYSCPATNECEITKRRKSC 180
Db 121 NSMPKRLCLVCGDIAGGYHGVASCEACAFPERKTIQGNIEYSCPATNECEITKRRKSC 180
Qy 181 QACRFPMKCLKVGMLKEGVRLDRVGRGOKYKRRIDENSPYLNPOQVOPAKKPYNKIVSH 240
Db 181 QACRFPMKCLKVGMLKEGVRLDRVGRGOKYKRRIDENSPYLNPOQVOPAKKPYNKIVSH 240
Qy 241 LVAEBEKTIVAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHIPGSTSLADQMSL 300
Db 241 LVAEBEKTIVAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHIPGSTSLADQMSL 300
Qy 301 LOSAMEIILIGVYVYSLSFEDELVYADYIMEDOSKLAGLIDINNAILQIVKYYKSMK 360
Db 301 LOSAMEIILIGVYVYSLSFEDELVYADYIMEDOSKLAGLIDINNAILQIVKYYKSMK 360
Qy 361 LEKEEFVTLKATIALANDSMHIEDVAVOKLODVLEHALQDYRAGOHMEDPRRAGKMLMT 420
Db 361 LEKEEFVTLKATIALANDSMHIEDVAVOKLODVLEHALQDYRAGOHMEDPRRAGKMLMT 420
Qy 421 LPLRLQSTKAVQHFYNIKLEGVPMHKLPLEMLBAKV 458
Db 421 LPLRLQSTKAVQHFYNIKLEGVPMHKLPLEMLBAKV 458
```

```
RESULT 2
US-10-355-218-1
/ Sequence 1, Application US/10355218
/ Publication No. US20040009558A1
/ GENERAL INFORMATION:
/ APPLICANT: MORAS, DINO
/ APPLICANT: RENAUD, JEAN-PAUL
/ APPLICANT: GRESCHIK, HOLGER
/ APPLICANT: WURTZ, JEAN-MARIE
/ TITLE OF INVENTION: THE LIGAND BINDING POCKET PEPTIDE FRAGMENT OF THE
/ TITLE OF INVENTION: ESTROGEN-RELATED RECEPTOR 3 (ERR3) AND USES THEREOF
/ FILE REFERENCE: 0510-1060
/ CURRENT APPLICATION NUMBER: US/10/355, 218
/ CURRENT FILING DATE: 2003-01-31
/ PRIOR APPLICATION NUMBER: 60/352, 551
/ PRIOR FILING DATE: 2002-01-31
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 458
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-355-218-1
```

```
Query Match 100.0%; Score 2388; DB 15; Length 458;
Best Local Similarity 100.0%; Pred. No. 2,4e-215;
Matches 458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MSVELCLPESFSLHYEEELLCRMNKKDRHIDSSCSFITEPSPASLTDSVNHSPGG 60
Db 1 MSVELCLPESFSLHYEEELLCRMNKKDRHIDSSCSFITEPSPASLTDSVNHSPGG 60
Qy 61 SSDASGYSSTWNGHNGDSDPLYPSPAPILGSSGVPKRLYDDCSSTIVEDPOTKCEYML 120
Db 61 SSDASGYSSTWNGHNGDSDPLYPSPAPILGSSGVPKRLYDDCSSTIVEDPOTKCEYML 120
Qy 121 NSMPKRLCLVCGDIAGGYHGVASCEACAFPERKTIQGNIEYSCPATNECEITKRRKSC 180
Db 121 NSMPKRLCLVCGDIAGGYHGVASCEACAFPERKTIQGNIEYSCPATNECEITKRRKSC 180
Qy 181 QACRFPMKCLKVGMLKEGVRLDRVGRGOKYKRRIDENSPYLNPOQVOPAKKPYNKIVSH 240
Db 181 QACRFPMKCLKVGMLKEGVRLDRVGRGOKYKRRIDENSPYLNPOQVOPAKKPYNKIVSH 240
Qy 241 LVAEBEKTIVAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHIPGSTSLADQMSL 300
```

```
Db 241 LVAEBEKTIVAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHIPGSTSLADQMSL 300
Qy 301 LOSAMEIILIGVYVYSLSFEDELVYADYIMEDOSKLAGLIDINNAILQIVKYYKSMK 360
Db 301 LOSAMEIILIGVYVYSLSFEDELVYADYIMEDOSKLAGLIDINNAILQIVKYYKSMK 360
Qy 361 LEKEEFVTLKATIALANDSMHIEDVAVOKLODVLEHALQDYRAGOHMEDPRRAGKMLMT 420
Db 361 LEKEEFVTLKATIALANDSMHIEDVAVOKLODVLEHALQDYRAGOHMEDPRRAGKMLMT 420
Qy 421 LPLRLQSTKAVQHFYNIKLEGVPMHKLPLEMLBAKV 458
Db 421 LPLRLQSTKAVQHFYNIKLEGVPMHKLPLEMLBAKV 458
```

```
RESULT 3
US-10-054-841-6
/ Sequence 6, Application US/10054841
/ Publication No. US20030119100A1
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Fang
/ TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
/ TITLE OF INVENTION: RECEPTOR PROTEINS
/ FILE REFERENCE: 19999YCA
/ CURRENT APPLICATION NUMBER: US/10/054, 841
/ PRIOR APPLICATION NUMBER: 09/487, 379
/ PRIOR FILING DATE: 2000-01-18
/ PRIOR APPLICATION NUMBER: 09/441, 000
/ PRIOR FILING DATE: 1998-08-26
/ PRIOR APPLICATION NUMBER: 60/078, 633
/ PRIOR FILING DATE: 1998-03-19
/ PRIOR APPLICATION NUMBER: 60/062, 902
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/057, 090
/ PRIOR FILING DATE: 1997-08-27
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PaetsSBQ for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 418
/ TYPE: PRT
/ ORGANISM: Human
US-10-054-841-6
```

```
Query Match 89.6%; Score 2139; DB 14; Length 418;
Best Local Similarity 100.0%; Pred. No. 5,4e-194;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MSVELCLPESFSLHYEEELLCRMNKKDRHIDSSCSFITEPSPASLTDSVNHSPGG 60
Db 1 MSVELCLPESFSLHYEEELLCRMNKKDRHIDSSCSFITEPSPASLTDSVNHSPGG 60
Qy 61 SSDASGYSSTWNGHNGDSDPLYPSPAPILGSSGVPKRLYDDCSSTIVEDPOTKCEYML 120
Db 61 SSDASGYSSTWNGHNGDSDPLYPSPAPILGSSGVPKRLYDDCSSTIVEDPOTKCEYML 120
Qy 121 NSMPKRLCLVCGDIAGGYHGVASCEACAFPERKTIQGNIEYSCPATNECEITKRRKSC 180
Db 121 NSMPKRLCLVCGDIAGGYHGVASCEACAFPERKTIQGNIEYSCPATNECEITKRRKSC 180
Qy 181 QACRFPMKCLKVGMLKEGVRLDRVGRGOKYKRRIDENSPYLNPOQVOPAKKPYNKIVSH 240
Db 181 QACRFPMKCLKVGMLKEGVRLDRVGRGOKYKRRIDENSPYLNPOQVOPAKKPYNKIVSH 240
Qy 241 LVAEBEKTIVAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHIPGSTSLADQMSL 300
Db 241 LVAEBEKTIVAMPDPTVPDSDIKALTTLCDLADRELVIIGMAKHIPGSTSLADQMSL 300
Qy 301 LOSAMEIILIGVYVYSLSFEDELVYADYIMEDOSKLAGLIDINNAILQIVKYYKSMK 360
Db 301 LOSAMEIILIGVYVYSLSFEDELVYADYIMEDOSKLAGLIDINNAILQIVKYYKSMK 360
```

QY 361 LEKEEFTVTLKAIALANSDSMHEIDVEAVQKLDVYHEALQDYEAQOME 409
 Db 361 LEKEEFTVTLKAIALANSDSMHEIDVEAVQKLDVYHEALQDYEAQOME 409

RESULT 4

US-10-054-841-2
 ; Sequence 2, Application US/10054841
 ; Publication No. US20030119100A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Fang
 ; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
 ; TITLE OF INVENTION: RECEPTOR PROTEINS
 ; FILE REFERENCE: 19999YCA
 ; CURRENT APPLICATION NUMBER: US/10/054,841
 ; CURRENT FILING DATE: 2002-01-23
 ; PRIOR APPLICATION NUMBER: 09/487,379
 ; PRIOR FILING DATE: 2000-01-18
 ; PRIOR APPLICATION NUMBER: 09/141,000
 ; PRIOR FILING DATE: 1998-08-26
 ; PRIOR APPLICATION NUMBER: 60/078,633
 ; PRIOR FILING DATE: 1998-03-19
 ; PRIOR APPLICATION NUMBER: 60/062,902
 ; PRIOR FILING DATE: 1997-10-21
 ; PRIOR APPLICATION NUMBER: 60/057,090
 ; PRIOR FILING DATE: 1997-08-27
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 500
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-054-841-2

Query Match 75.3%; Score 1799; DB 14; Length 500;
 Best Local Similarity 78.1%; Pred. No. 6.5e-160;
 Matches 339; Conservative 50; Mismatches 43; Indels 2; Gaps 2;

QY 24 MSNKDRHIDSSCSFITEPSSPASLTDVYNHSPGSSSDASGSYSTNNGHONGLDSP 83
 Db 1 MSSEDRHIDSSCSFITEPSSPASLTDVYNHSPGSSSDASGSYSTNNGHONGLDSP 60
 84 LVPSAPLIGSGGPRKLYDCCSTIVEDPOTKCEYMLNSMPKRLCLVCGDIAAGYHGA 143
 Db 61 MPAGAG-LGGT-PCRKSEYEDCAGIMEDSAIKCEYMLNAPKRLCLVCGDIAAGYHGA 118
 QY 144 SCEACAFKFRITIGNIEYSCPATNECEITKRRKSCQACRFMKCLKVGMKGVRLDRV 203
 Db 119 SCEACAFKFRITIGNIEYSCPATNECEITKRRKSCQACRFMKCLKVGMKGVRLDRV 178
 QY 204 RGRQKRRKRLIDANSFYLNPOLVQPAKKRYNKIVSHLLVAEPKTIYAMPDPVYPSDIX 263
 Db 179 RGRQKRRKRLIDANSFYLNPOLVQPAKKRYNKIVSHLLVAEPKTIYAMPDPVYPSDIX 238
 QY 264 ALTTLCDLADRELVIITIGMAKHIPGFSTLSLADOMSLQSAWMEILILGVVYSLSFEDE 323
 Db 239 ALTTLCDLADRELVIITIGMAKHIPGFSTLSLADOMSLQSAWMEILILGVVYSLSFEDE 298
 QY 324 LVYADYIMDEDSKLAGLIDLNNAILOVYKYSKMLEKEEFTVTLKAIALANSDSMHE 383
 Db 299 LVYADYIMDEHSRLAGLLELYRAILQIVRRYKVKLVEKEEFTVTLKAIALANSDSMHE 358
 QY 384 DVAEAVQKLDVYHEALQDYEAQOMEDPRAQKMLMTPLIRCTSTYAAVOHFYNITKEGK 443
 Db 359 DVAEAVQKLDVYHEALQDYELQSHHEBPWRTGKLLTLPLIRCTYAAVOHFYSVVKLGK 418
 QY 444 VPMHKLFLMULEAK 457
 Db 419 VPMHKLFLMULEAK 432

RESULT 5
 US-10-278-481-19

Sequence 19, Application US/10278481
 Publication No. US20030113803A1
 GENERAL INFORMATION:
 APPLICANT: KARO BIO AB
 TITLE OF INVENTION: Orphan receptor
 NUMBER OF SEQUENCES: 19
 STREET: c/o Center for Biotechnology and Department
 of Medical Nutrition, Karolinska Nutrition, Karolinska
 Institute
 CITY: Huddinge
 COUNTRY: Sweden
 ZIP: S-14186
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/278,481
 FILING DATE: 23-Oct-2002
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/333,057
 FILING DATE: <Unknown>
 APPLICATION NUMBER: 08/836,620
 FILING DATE: <Unknown>
 APPLICATION NUMBER: GB 9518272.1
 FILING DATE: 08-SEP-1995
 APPLICATION NUMBER: GB 9605550.4
 FILING DATE: 15-MAR-1996
 APPLICATION NUMBER: GB 9607532.0
 FILING DATE: 11-APR-1996
 APPLICATION NUMBER: GB 9609576.5
 FILING DATE: 08-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: GUSTAFSSON, Jan-Ake
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 431 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 SEQUENCE DESCRIPTION: SEQ ID NO: 19:
 US-10-278-481-19

Query Match 74.3%; Score 1775; DB 14; Length 431;
 Best Local Similarity 76.4%; Pred. No. 9.3e-158;
 Matches 331; Conservative 51; Mismatches 49; Indels 2; Gaps 1;

QY 24 MSNKDRHIDSSCSFITEPSSPASLTDVYNHSPGSSSDASGSYSTNNGHONGLDSP 83
 Db 1 MSSEDRHIDSSCSFITEPSSPASLTDVYNHSPGSSSDASGSYSTNNGHONGLDSP 60
 84 LVPSAPLIGSGGPRKLYDCCSTIVEDPOTKCEYMLNSMPKRLCLVCGDIAAGYHGA 143
 Db 61 MPAGAGL--GPNCRKSEYEDCTGIMEDSAIKCEYMLNAPKRLCLVCGDIAAGYHGA 118
 QY 144 SCEACAFKFRITIGNIEYSCPATNECEITKRRKSCQACRFMKCLKVGMKGVRLDRV 203
 Db 119 SCEACAFKFRITIGNIEYSCPATNECEITKRRKSCQACRFMKCLKVGMKGVRLDRV 178
 QY 204 RGRQKRRKRLIDANSFYLNPOLVQPAKKRYNKIVSHLLVAEPKTIYAMPDPVYPSDIX 263
 Db 179 RGRQKRRKRLIDANSFYLNPOLVQPAKKRYNKIVSHLLVAEPKTIYAMPDPVYPSDIX 238
 QY 264 ALTTLCDLADRELVIITIGMAKHIPGFSTLSLADOMSLQSAWMEILILGVVYSLSFEDE 323
 Db 239 ALTTLCDLADRELVIITIGMAKHIPGFSTLSLADOMSLQSAWMEILILGVVYSLSFEDE 298
 QY 324 LVYADYIMDEDSKLAGLIDLNNAILOVYKYSKMLEKEEFTVTLKAIALANSDSMHE 383
 Db 299 LVYADYIMDEHSRLAGLLELYRAILQIVRRYKVKLVEKEEFTVTLKAIALANSDSMHE 358

QY 384 DVEAVOKLQDVHLEALQDYEAQGHMEDPRRAGKMLTLLPRLROTSTKAVOHFNKLEK 443
DB 359 NIEAVOKLQDVLHEALQDYELSQHREBPRRAGKMLTLLPRLROTAKAVOHFNYSVKLOK 418
QY 444 VPMHKLFLMLEA 456
DB 419 VPMHKLFLMLEA 431

RESULT 6

US-10-297-653-1
Sequence 1, Application US/10297653
Publication No. US20030215829A1
GENERAL INFORMATION:
APPLICANT: INCTB GENOMICS, INC.
APPLICANT: CHINN, Anna M.
APPLICANT: LU, Yan
APPLICANT: TRIBOULEY, Catherine M.
TITLE OF INVENTION: NUCLEAR HORMONE RECEPTORS
FILE REFERENCE: PI-0119 USN
CURRENT APPLICATION NUMBER: US/10/297,653
PRIOR FILING DATE: 2002-12-06
PRIOR APPLICATION NUMBER: US 60/209,863
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 486
TYPE: PRF
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030215829A1 7475568CD1
US-10-297-653-1

Query Match 52.7%; Score 1259; DB 15; Length 486;
Best Local Similarity 54.9%; Pred. No. 3.6e-109;
Matches 256; Conservative 67; Mismatches 79; Indels 64; Gaps 8;

QY 22 CSMNKRHHIDSSCS-----FKTEPSSPASITDSVNHSPGSSD-----A 64
DB 51 CPTALPEPQVTAMSQVVGIEPLIYKABPASP-----DSPKSSSETEPEPVALA 101
QY 65 SG-SYSTMNGHONGDSDPLVPSPAPILGSGSPVKRYLDDCSSTIYEDQTCENYLNEM 123
DB 102 PGPAPTRCLPGHKEEDGE-----GAGP-----GQGGGKVLSSU 137
QY 124 PRLCLVCGDIASGYHYGASCEACAFKRTIQGNIEXSCPATNECEITKRRKSCQAC 183
DB 138 PRLCLVCGDVASGYHYGASCEACAFKRTIQGSIEXSCPASNECEITKRRKSCQAC 197
QY 184 RMKCLKVAMKEGVRLDVRGGRQKRYKRIDAENSPYINPOLVOP-----AKRY 234
DB 198 RETKCLRVMLEKGVRLDVRGGRQKRYKRPEVDPLPFGPPAGPLAVAGGRKTAAPV 257
QY 235 NKTIVSHLVAPEPKIYAMPDPVPPDSIDIKATLTLCDLADRELVIITGMAGHIFGFTLSU 294
DB 258 NKTIVSHLVAPEPKIYAMPDPDPAGDGHLPVATLCLDFREIYVITISMAKSLPFGFTLSU 317
QY 295 ADOMSLQSAWMEIILIGVYRSLSPEDELVYADYIMEDOSKLAGLIDNNAILQVYK 354
DB 318 SPQMSVILQSVWMEVILVGVAGRSILPQDELAPAFEDLVLDBEGARAAGLIGLALTLQVYR 377
QY 355 KYKSMKLEKEEFTYTKAIALANDSMHIEVEAVOKLQDVHLEALQDYEAQO---HMD 410
DB 378 RLQALRLEREVEYLLKALALANDSVHIEDEAVEQRLRAHLHEALTEYARAGPAGGAB 437
QY 411 PRRAGKMLTLLPRLROTSTKAVOHFNKLEKGVPMHKLFLMLEA 456
DB 438 RRRAGKMLTLLPRLROTAKAVLAHFYGVKLEKGVPMHKLFLMLEA 483

RESULT 7

US-10-211-239-1
Sequence 1, Application US/10211239
Publication No. US20030103965A1
GENERAL INFORMATION:
APPLICANT: Jung, Birgit
APPLICANT: Kraut, No. US20030103965A1bert
APPLICANT: Mueller, Stefan
TITLE OF INVENTION: Method for Identifying Substances Which Positively
TITLE OF INVENTION: Influence Inflammatory Conditions
FILE REFERENCE: 0652.2340001
CURRENT APPLICATION NUMBER: US/10/211,239
PRIOR FILING DATE: 2002-08-05
PRIOR APPLICATION NUMBER: US 60/315,775
PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: EP 0119003.0
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 519
TYPE: PRF
ORGANISM: Homo Sapiens
US-10-211-239-1

Query Match 52.7%; Score 1258.5; DB 14; Length 519;
Best Local Similarity 55.1%; Pred. No. 4.4e-109;
Matches 256; Conservative 67; Mismatches 79; Indels 63; Gaps 8;

QY 22 CSMNKRHHIDSSCS-----FKTEPSSPASITDSVNHSPGSSD-----A 64
DB 85 CPTALPEPQVTAMSQVVGIEPLIYKABPASP-----DSPKSSSETEPEPVALA 135
QY 65 SG-SYSTMNGHONGDSDPLVPSPAPILGSGSPVKRYLDDCSSTIYEDQTCENYLNEM 123
DB 136 PGPAPTRCLPGHKEEDGE-----GAGP-----GQGGGKVLSSU 171
QY 124 PRLCLVCGDIASGYHYGASCEACAFKRTIQGNIEXSCPATNECEITKRRKSCQAC 183
DB 172 PRLCLVCGDVASGYHYGASCEACAFKRTIQGSIEXSCPASNECEITKRRKSCQAC 231
QY 184 RMKCLKVAMKEGVRLDVRGGRQKRYKRIDAENSPYINPOLVOP-----AKRY 235
DB 232 RETKCLRVMLEKGVRLDVRGGRQKRYKRPEVDPLPFGPPAGPLAVAGGRKTAAPV 291
QY 236 KIVSHLVAPEPKIYAMPDPVPPDSIDIKATLTLCDLADRELVIITGMAGHIFGFTLSU 295
DB 292 ALVSHLVAPEPKIYAMPDPAGDGHLPVATLCLDFREIYVITISMAKSLPFGFTLSU 351
QY 296 DQMSLQSAWMEIILIGVYRSLSPEDELVYADYIMEDOSKLAGLIDNNAILQVYK 355
DB 352 DQMSVILQSVWMEVILVGVAGRSILPQDELAPAFEDLVLDBEGARAAGLIGLALTLQVYR 411
QY 356 YKSMKLEKEEFTYTKAIALANDSMHIEVEAVOKLQDVHLEALQDYEAQO---HMD 411
DB 412 LQALRLEREVEYLLKALALANDSVHIEDEAVEQRLRAHLHEALTEYARAGPAGGAB 471
QY 412 RRRAGKMLTLLPRLROTSTKAVOHFNKLEKGVPMHKLFLMLEA 456
DB 472 RRRAGKMLTLLPRLROTAKAVLAHFYGVKLEKGVPMHKLFLMLEA 516

RESULT 8

US-10-278-481-18
Sequence 18, Application US/10278481
Publication No. US20030113803A1
GENERAL INFORMATION:
APPLICANT: KARO BIO AB
TITLE OF INVENTION: Orphan receptor
NUMBER OF SEQUENCES: 19
STREET: c/o Center for Biotechnology and Department
of Medical Nutrition, Karolinska Nutrition, Karolinska
Institute
CITY: Huddinge


```
/ COUNTRY: Sweden
/ ZIP: S-14186
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
/ CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/10/278,481
/   FILING DATE: 23-Oct-2002
/ PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: US/09/333,057
/   FILING DATE: <Unknown>
/   APPLICATION NUMBER: 08/836,620
/   FILING DATE: <Unknown>
/   APPLICATION NUMBER: GB 9518272.1
/   FILING DATE: 08-SEP-1995
/   APPLICATION NUMBER: GB 9605550.4
/   FILING DATE: 15-MAR-1996
/   APPLICATION NUMBER: GB 9607532.0
/   FILING DATE: 11-APR-1996
/   APPLICATION NUMBER: GB 9609576.5
/   FILING DATE: 08-MAY-1996
/ ATTORNEY/AGENT INFORMATION:
/   NAME: GUSTAFSSON, Jan-Ake
/   INFORMATION FOR SEQ ID NO: 18:
/     SEQUENCE CHARACTERISTICS:
/       LENGTH: 518 amino acids
/       TYPE: amino acid
/       TOPOLOGY: linear
/ ORIGINAL SOURCE:
/   ORGANISM: Homo sapiens
/   SEQUENCE DESCRIPTION: SEQ ID NO: 18:
/ US-10-278-481-18
/
/ Query Match      50.8%; Score 1212.5; DB 14; Length 518;
/ Best Local Similarity 53.5%; Pred. No. 9.3e-105;
/ Matches 250; Conservative 67; Mismatches 85; Indels 65; Gaps 9;
/
/ QY 22 CRKSNCRHIDSSGSS-----FITFSPSPASLIDSVNHHSPGGSSD-----A 64
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   85 CPTALPPOVTSAMSSQVGIETPLIKAEPAAP-----DSPKSSSTETETPPVALA 135
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 65 SG-SYSTMNGHONGDSPPLYSAPILGSSGPYRKIYDCCSTIYVDPQTKCRYMNSM 123
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   136 PGAPAPRCLPGHKEBEDGE-----GAGP-----GEOGGGKLVLSL 171
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 124 PKRLCTVCGDIASGYHYGVASCEACAFKRTIQGNIYESCPATNECEITKRRKSCQAC 183
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   172 PKRLCTVCGDVASGYHYGVASCEACAFKRTIQGSIIEYSCPSANECEITKRRKSCQAC 231
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 184 RFKKCKLVGMLKEGVLDPRGRGRQYKRRIDANSPYLNPLQVQ-----AKPY 234
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   232 RFTKCKRVGMLKEGVLDPRGRGRQYKRRIDANSPYLNPLQVQ-----AKPY 291
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 235 NKIVSHLVAEPEKIYAMPDPPTVPDSIDKALTLCDLADDELYIIGMAKHITGSETLSL 294
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   292 NLIVSHLVAEPEKIYAMPDPDPDGLPAVATLCPDDEIYVTTISMAKSIIPGFSLSL 351
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 295 ADOMSLIGSAWEIILIGVYRSISFEDELIVADYIMEDOSKLAGLIDLNNAIIQLVYK 354
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   352 SDQMSTVIGSVMMEVILGVAGRSPLQDELAFADVLILBEGRAAGLGLGALLQLVYK 411
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   355 KTKSMLEKEEPTTKAIALNSDSNHEIDVEAV-OKLQDVHEALQDYBAGQ---HME 409
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   412 RIQALULEEVEEYVLAALANSDSVHIDDEPLTMSCKLHEALLFEYBAGAGGCGGA 471
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/ QY 410 DPRRAGKMLTPLLQSTKAQHFYNTLBEKGVMHKLPLEMEEA 456
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/   472 FERRRAGRILLTLPLRQITAGKYLAAHYGYVLEKSKVMHKLPLEMEEA 518
/
/ RESULT 9
```

```
US-10-093-463-202
/ Sequence 202, Application US/10093463
/ Publication No. US20030208039A1
/ GENERAL INFORMATION:
/   APPLICANT: Padigaru, Muralidhara
/   APPLICANT: Shenoy, Suresh
/   APPLICANT: Kexuda, Ramesh
/   APPLICANT: Gusev, Vladimir
/   APPLICANT: Pochart, Pascal
/   APPLICANT: Zhong, Mei
/   APPLICANT: Rastelli, Luca
/   APPLICANT: Mezes, Peter
/   APPLICANT: Smithson, Glenda
/   APPLICANT: Guo, Xiaojia
/   APPLICANT: Gerlach, Valerie
/   APPLICANT: Casman, Stacie
/   APPLICANT: Boldog, Ferenc
/   APPLICANT: Li, Li
/   APPLICANT: Zerhusen, Bryan
/   APPLICANT: Tchenev, Velizar
/   APPLICANT: Gangolli, Esma
/   APPLICANT: Vernet, Corinne
/   APPLICANT: Pena, Carol
/   APPLICANT: Burgess, Catherine
/   APPLICANT: Liu, Xiaohong
/   APPLICANT: Spytek, Kimberly
/   APPLICANT: Gorman, Linda
/   APPLICANT: Spaderna, Steven
/   APPLICANT: Voss, Edward
/   APPLICANT: Malyankar, Uriel
/   APPLICANT: Anderson, David
/   APPLICANT: Patturajan, Meera
/   APPLICANT: Miller, Charles
/   APPLICANT: Taupier, Raymond J. Jr.
/ TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypeptide
/ FILE REFERENCE: Encoding The Antigens, and Methods of Use.
/ CURRENT APPLICATION NUMBER: US/10/093,463
/ PRIOR APPLICATION NUMBER: 2002-06-24
/ PRIOR FILING DATE: 2001-04-14
/ PRIOR APPLICATION NUMBER: 60/358,092
/ PRIOR FILING DATE: 2001-12-03
/ PRIOR APPLICATION NUMBER: 60/274,281
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/274,101
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/325,681
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: 60/304,354
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/279,995
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 60/294,899
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 60/287,424
/ PRIOR FILING DATE: 2001-04-30
/ PRIOR APPLICATION NUMBER: 60/299,027
/ PRIOR FILING DATE: 2001-06-18
/ PRIOR APPLICATION NUMBER: 60/309,198
/ PRIOR FILING DATE: 2001-07-31
/ PRIOR APPLICATION NUMBER: 60/281,194
/ PRIOR FILING DATE: 2001-04-04
/ PRIOR APPLICATION NUMBER: 60/274,194
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/274,849
/ PRIOR FILING DATE: 2001-03-09
/ PRIOR APPLICATION NUMBER: 60/330,380
/ PRIOR FILING DATE: 2001-10-18
/ PRIOR APPLICATION NUMBER: 60/275,235
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: 60/288,342
/ PRIOR FILING DATE: 2001-05-03
```

PRIOR APPLICATION NUMBER: 60/275,578

PRIOR FILING DATE: 2001-03-13

NUMBER OF SEQ ID NOS: 370

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 202

LENGTH: 516

TYPE: PRT

ORGANISM: Homo sapiens

US-10-093-463-202

Query Match 44.9%; Score 1073; DB 15; Length 516;
Best Local Similarity 50.4%; Pred. No. 1,2e-91;
Matches 226; Conservative 68; Mismatches 98; Indels 56; Gaps 9;

```

QY 31 IDSSCSFIKTPSSPASLTDSVNHSPGSSD-----ASG-STYSSTMGHQGLD 80
DB 100 VGSFPLQYIMBAPAP-----DSPKGSSEETEPVALAPGPAPHTCLPGKEED 150
QY 81 SPPLYSAPILGSGPVAKLYDSCSTIVEDPQTCFYMLNMPKRLCYVCGDIASGYHY 140
DB 151 GE-----GAGP-----GEGGGKVLVSLSKRLCLVCGDVASGYHC 186
QY 141 GVASCEACAFKFKRTIQGNIETSCPATNECEITKRRKSCQACREMKCLKVMKEGYVL 200
DB 187 GVSCEDECKAFKRTIQGSMETSCASNECEITKRRKSCQACRETKSLRLHQEPAGARL 246
QY 201 DRVGRGRQYKRRIRIENSFY-----LNPQVQPAK--PYNKIVSHLVAEPEKIYAM 252
DB 247 DRVGRGRQYKRRIRIENSFY-----LNPQVQPAK--PYNKIVSHLVAEPEKIYAM 305
QY 253 PPTVPDSDIKALTTLCDADELIVYIIGMAKHIFGFTSLIADQMSLQASAMETILIG 312
DB 306 PPPAGDGGHLPVATICDLFDEIIVITISMAKSIPEFSSLSIDQSVLVQRYMVEVLVPG 365
QY 313 VYVRSISFEDELIVYADYIMDEDOQSLAGLIDNNAILQVKKYKSMLEKEEFYTLKAI 372
DB 366 VAGRSIPLQDELAFARDLVDERGARAAGIGELGALLQVRLQSLRERGGYVLLKAL 425
QY 373 ALANSDSMIEDEVAOVKIQDVYHEALQDYEAQO---HMEDPRRAGKMLTLPILRQTS 428
DB 426 ALANSDSVIEDEVAEVAQLEAPHEALIEEAGRAQTGGGAERPRRGLTLPILHQTA 485
QY 429 TKAVQHFNKIKLEGKVMKHLFLEMLEA 456
DB 486 GKVLAHFYGVKLKGVPMKHLFLEMLEA 513

```

RESULT 10

US-10-355-218-8

Sequence 8, Application US/10355218

Publication No. US20040009558A1

GENERAL INFORMATION:

APPLICANT: MORAS, DINO

APPLICANT: RENAUD, JEAN-PAUL

APPLICANT: GRESCHIK, HOLGER

APPLICANT: WURTZ, JEAN-MARIE

TITLE OF INVENTION: THE LIGAND BINDING POCKET PEPTIDE FRAGMENT OF THE

TITLE OF INVENTION: ESTROGEN-RELATED RECEPTOR 3 (ERR3) AND USES THEREOF

FILE REFERENCE: 0510-1060

CURRENT APPLICATION NUMBER: US/10/355,218

CURRENT FILING DATE: 2003-01-31

PRIOR APPLICATION NUMBER: 60/352,551

PRIOR FILING DATE: 2002-01-31

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 8

LENGTH: 230

TYPE: PRT

ORGANISM: Homo sapiens

US-10-355-218-8

Query Match 39.7%; Score 947; DB 15; Length 230;

Best Local Similarity 79.4%; Pred. No. 2.4e-80;

Matches 181; Conservative 29; Mismatches 18; Indels 0; Gaps 0;

```

QY 230 AKRPYNKIVSHLVAEPEKIYAMPDPTVPDSDIKALTTLCDADELIVYIIGMAKHIFG 289
DB 1 AKRPYNKIVSHLVAEPEKIYAMPDPTVPDSDIKALTTLCDADELIVYIIGMAKHIFG 60
QY 290 STLSDQMSLQASAMMEIILIGVYRSLSFEDELIVYADYIMDEDOQSLAGLIDNNAIL 349
DB 61 STLSDQMSLQASAMMEIILIGVYRSLSFEDELIVYADYIMDEDOQSLAGLIDNNAIL 120
QY 350 LQVYKYSKMLEKEEFYTLKAIALANSDSMIEDEVAOVKIQDVYHEALQDYEAQO---HM 409
DB 121 LQVYKYSKMLEKEEFYTLKAIALANSDSMIEDEVAOVKIQDVYHEALQDYEAQO---HM 180
QY 410 DRRAGKMLTLPILRQTSKAVQHFYNIKLEGKVMKHLFLEMLEAK 457
DB 181 EPRMTRGLTLPILRQTSKAVQHFYNIKLEGKVMKHLFLEMLEAK 228

```

RESULT 11

US-10-355-218-9

Sequence 9, Application US/10355218

Publication No. US20040009558A1

GENERAL INFORMATION:

APPLICANT: MORAS, DINO

APPLICANT: RENAUD, JEAN-PAUL

APPLICANT: GRESCHIK, HOLGER

APPLICANT: WURTZ, JEAN-MARIE

TITLE OF INVENTION: THE LIGAND BINDING POCKET PEPTIDE FRAGMENT OF THE

TITLE OF INVENTION: ESTROGEN-RELATED RECEPTOR 3 (ERR3) AND USES THEREOF

FILE REFERENCE: 0510-1060

CURRENT APPLICATION NUMBER: US/10/355,218

CURRENT FILING DATE: 2003-01-31

PRIOR APPLICATION NUMBER: 60/352,551

PRIOR FILING DATE: 2002-01-31

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9

LENGTH: 234

TYPE: PRT

ORGANISM: Homo sapiens

US-10-355-218-9

Query Match 31.6%; Score 755; DB 15; Length 234;

Best Local Similarity 62.7%; Pred. No. 2.8e-62; Indels 4; Gaps 1;

Matches 143; Conservative 43; Mismatches 38; Indels 4; Gaps 1;

```

QY 233 PYNKIVSHLVAEPEKIYAMPDPTVPDSDIKALTTLCDADELIVYIIGMAKHIFG 292
DB 4 PYNKIVSHLVAEPEKIYAMPDPTVPDSDIKALTTLCDADELIVYIIGMAKHIFG 63
QY 293 STLSDQMSLQASAMMEIILIGVYRSLSFEDELIVYADYIMDEDOQSLAGLIDNNAILQ 352
DB 64 STLSDQMSLQASAMMEIILIGVYRSLSFEDELIVYADYIMDEDOQSLAGLIDNNAILQ 123
QY 353 VYKYSKMLEKEEFYTLKAIALANSDSMIEDEVAOVKIQDVYHEALQDYEAQO---HM 408
DB 124 VYKYSKMLEKEEFYTLKAIALANSDSMIEDEVAOVKIQDVYHEALQDYEAQO---HM 183
QY 409 EDRPRRAGKMLTLPILRQTSKAVQHFYNIKLEGKVMKHLFLEMLEA 456
DB 184 EDRPRRAGKMLTLPILRQTSKAVQHFYNIKLEGKVMKHLFLEMLEA 231

```

RESULT 12

US-10-450-484A-1

Sequence 1, Application US/10450484A

Publication No. US20040121355A1

GENERAL INFORMATION:

APPLICANT: Sumitomo Chemical Company Limited

TITLE OF INVENTION: Estrogen Receptor Genes and Utilization Thereof

FILE REFERENCE: 600630-5US (560178)

CURRENT APPLICATION NUMBER: US/10/450,484A

; CURRENT FILING DATE: 2004-02-09
 ; PRIOR APPLICATION NUMBER: JP 2000/379993
 ; PRIOR FILING DATE: 2000-12-14
 ; PRIOR APPLICATION NUMBER: PCT/JP01/09996
 ; PRIOR FILING DATE: 2001-11-15
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 587
 ; TYPE: PRT
 ; ORGANISM: Caiman crocodilus
 ; US-10-450-484A-1

Query Match 30.4%; Score 726; DB 16; Length 587;
 Best Local Similarity 38.2%; Pred. No. 6,1e-59;
 Matches 175; Conservative 80; Mismatches 143; Indels 60; Gaps 14;

QY 38 FIKTEBSSPALTDVSNHSPG-----SSDASGYSSTNGHONGDPPPLYSAPILGG 93
 Db 104 FLQTAF-----QLSPFVHHHQVYVYLENDQSG-----GMRBASSTYRPSADSRHQ 153
 QY 94 SGVVRKLYDCCSTIYEDPQTKCEYMLNSMPK-RLCLVCGDIASGYHYGVAACEACKAF 152
 Db 154 SGRER-----MSSST-----SEKASLSWESTKETREYCAVCDYASGYHYGVAACEACKAF 203
 QY 153 KRTIOGNIEXCPATNCEITKRRKSCQACRFMKCLXGMLKEGVRLDVRGGRQ-KYK 211
 Db 204 KRSLQGHNDYMCPTAINQCTIDKRRKSCQACRLKCYEVAMGKGLRKDR-RGRMKLKQK 262
 QY 212 RR--IDAENSP-----YLNPOLVOPAKK-----PYNKIVSHLVAPEKIV 250
 Db 263 RQEBQDARNGEYATAMERTPLMTSPLYVHKTKKSPALSTABQVNSALLREPEPIV 322
 QY 251 AMDPVPYSDIALTLTCLDLADRELYVITGMAKHIDFSTLSLADQMSLLQSAMMELI 310
 Db 323 SEYDEPKRPFVEASMTLTLTLADRELYVITGMAKHIDFSTLSLADQMSLLQSAMMELI 382
 QY 311 LGVYVRSLSFEDLVYADYIMDEDSQ-LAGLLDNLNALLQVKKYKSKLEKEPEVYL 369
 Db 383 IGVVMSMEHPGKLPAENLLDRNOQKCEGAVEIFDMLATPARRMMNLQGEFEVCL 442
 QY 370 KATALAN-----SDSMIEDVEAVOKLQVLYHEALQDYEAQOHM---EDPRRAGKL 418
 Db 443 KSIILNSGYVTFPLSTLSLEKDYIHRVLDKITITLILHMAKSGLSLQOQHRRLAQL 502
 QY 419 MTLPLRQSTKAVQHPYNIKLEGVPMKFLPEMLEA 456
 Db 503 LILSHIRHMSNKGMEHLVNMKCKVVPYLDLLEMLDA 540

RESULT 13
 US-10-278-481-16
 ; Sequence 16, Application US/10278481
 ; Publication No. US20030113803A1
 ; GENERAL INFORMATION:

; APPLICANT: KARO BIO AB
 ; TITLE OF INVENTION: Orphan receptor
 ; NUMBER OF SEQUENCES: 19
 ; STREET: c/o Center for Biotechnology and Department
 ; of Central Nutrition, Karolinska Nutrition, Karolinska
 ; Institute

; CITY: Huddinge
 ; COUNTRY: Sweden
 ; ZIP: S-14186

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/278,481
 ; FILING DATE: 23-Oct-2002
 ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/333,057
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: 08/836,620
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: GB 9518272.1
 ; FILING DATE: 08-SEP-1995
 ; APPLICATION NUMBER: GB 9605550.4
 ; FILING DATE: 15-MAR-1996
 ; APPLICATION NUMBER: GB 9607532.0
 ; FILING DATE: 11-APR-1996
 ; APPLICATION NUMBER: GB 9609576.5
 ; FILING DATE: 08-MAY-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: GUSTAFSSON, Jan-Ake
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 596 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; ORIGINAL SOURCE:
 ; ORGANISM: Rattus rattus
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
 ; US-10-278-481-16

Query Match 30.0%; Score 716.5; DB 14; Length 596;
 Best Local Similarity 39.2%; Pred. No. 4.9e-58;
 Matches 168; Conservative 82; Mismatches 128; Indels 51; Gaps 13;

QY 73 NGHONGLDSPRY-----PSAPILGGSGP--VAKLYDDCSSTVEPQTCF---YMLN 121
 Db 129 HGHQ-----VPYLENPPSAVAVDTGPPAFYRNSDNRQNGERLSSESSEKMMIMES 183
 QY 122 SWPKRLCLVCGDIASGYHYGVAACEACKAFKRTIOGNIEXCPATNCEITRRRRSCQ 181
 Db 184 AKETRCVAVNDYASGYHYGVAACEACKAFKRSIQGHNDYMCPTAINQCTIDKRRKSCQ 243
 QY 182 ACPFMKCLVGMLEKGVRLDVRGGRQ-KYK-RIDAEN-----SPYLNPOL 226
 Db 244 ACRLRCYEGVMKGGIRKDR-RGGRMLKHKRQRDLEGNENMGSGDMRAAMLWESPLV 302
 QY 227 VOPAKK-----PYNKIVSHLVAPEKIVAMPDPTVPSPDIKALTTLCDLADRELYV 279
 Db 303 IKHTKNSPALSTPAQVMSALIDAEPPILYSYDSRPFSEASMMGLTLNLADRELYVH 362
 QY 280 IGMAKHIDFSTLSLADQMSLLQSAMMELIIGVYVRSLSFEDLVYADYIMDEDSQK- 338
 Db 363 IMNAKVPKPGDNLNHDVHLLFCAMLEILMIGLVKRSMEHPGKLPAENLLDRNOQKC 422
 QY 339 LAGLLDNLNALLQVKKYKSKLEKEPEVTLKATALAN-----SDSMIEDVEAVOK 390
 Db 423 VEGMVEIFDMLATSSRFMMNLQGEFEVCLSIILNSGYVTFPLSTLSLEKDHHR 482
 QY 391 LDVLYHEALQDY--EAGQEMEDP-RRAGKMLTLPLRQSTKAVQHPYNIKLEGVPMH 447
 Db 483 VLDKINDYILHMAKAGTLTQOQHRRLAQLLILSHIRHMSNKGMEHLVNMKCKVVPY 542
 QY 448 KLEMLEEA 456
 Db 543 DILLEMLDA 551

RESULT 14
 US-10-451-768A-1

; Sequence 1, Application US/10451768A
 ; Publication No. US20040115765A1

; GENERAL INFORMATION:

; APPLICANT: Sumitomo Chemical Company Limited
 ; TITLE OF INVENTION: Estrogen receptor gene and its use
 ; FILE REFERENCE: 560163

; CURRENT APPLICATION NUMBER: US/10/451,768A

; CURRENT FILING DATE: 2003-12-24
 ; PRIOR APPLICATION NUMBER: JP 2000/392262
 ; PRIOR FILING DATE: 2000-12-25

NUMBER OF SEQ ID NOS: 27
 SEQ ID NO 1
 LENGTH: 581
 TYPE: PRT
 ORGANISM: Chemidophorus uniparens
 US-10-451-768A-1

Query Match 29.9%; Score 715; DB 16; Length 581;
 Best Local Similarity 34.9%; Pred. No. 6.5e-58;
 Matches 164; Conservative 80; Mismatches 133; Indels 93; Gaps 10;

QY 30 HIDSSCSFIKTEPSPASLT-----DSVNHSPGSSSDASGSSSTMGHNGML 79
 Db 117 HHNOQVPEYLENEBSSAMREAFPTAFYRPGSENNHGGRASSNSKSL----- 166
 QY 80 DSPPLPSAPILGGSPVKLYDCSSTIVBDPQTCRYMNSMKRILCVCGDIASGYH 139
 Db 167 -----MESTKETRYCAVCNDYASGYH 187
 QY 140 YGVASCECKAFKRTTIOGNIIEYSCPATNECEITKRRRKSQACRFEMCKIKVMKEGVR 199
 Db 188 YGVASCECKAFKRSIQGHNDYMCPATNQCTIDKRRKSQACRLKCYEGMMKGGIR 247
 QY 200 LDRVAGRGQ-KYKRRID-----AENSP-YLNPOLVOPAKK-----PYNKIV 238
 Db 248 KDR-RGGEMLKHKRRDLDGNNAVAVTEARNTILMPSPIMIKSKNSPALSLTAEGMV 306
 QY 239 SHLVAPEKTYAMPDPVPSDIALTTLCDLADRELVTIIGMAKHIPGSESTSLADQM 298
 Db 367 SALDLAEPPIVSEYDPSFSEASVMTLLTNLADRELVTIMTAKRYGFDLIDHGV 366
 QY 299 SILGSAMWEILLGVYRSLSFEDELVYADYIMEDQSK-LAGLIDLNNAILOLVKKYK 357
 Db 367 HLECAWELITMIGLWLSLHPGKILFAPNLLDRSQCMVEGVEIFDMILATSSRR 426
 QY 358 SMKLEKEEVTLKATALAN-----SPSMHIEDVEAVOKLODYLHEALQDYEAGQHN- 408
 Db 427 MNNGEGHEVVCILINSIGYITLSLRLSEKEHHRVLDKTTDLTHLMKSGLS 486
 QY 409 -EDPRRAGKMLTLPRLRQSTKAVQHPFYNIKLGKVMHKLFLBMLEA 456
 Db 487 LQQHRRILAQLMLSHIRMSNKGMEHLYNNKCKNVPLDYDLLEMLDA 536

RESULT 15

US-10-052-092-31
 ; Sequence 31, Application US/10052092
 ; Publication No. US20030027778A1

GENERAL INFORMATION:
 APPLICANT: Fugua, Suzanne
 APPLICANT: Allied, D.
 APPLICANT: Hopp, Torsten A.
 APPLICANT: O'Connell, Peter
 TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapeuti
 FILE REFERENCE: P02102US2
 CURRENT APPLICATION NUMBER: US/10/052,092
 CURRENT FILING DATE: 2002-01-18
 PRIOR APPLICATION NUMBER: US 60/262,990
 PRIOR FILING DATE: 2001-01-19
 PRIOR APPLICATION NUMBER: US 60/304,018
 PRIOR FILING DATE: 2001-07-09
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 31
 LENGTH: 595
 TYPE: PRT
 ORGANISM: pig
 US-10-052-092-31

Query Match 29.9%; Score 713.5; DB 14; Length 595;
 Best Local Similarity 42.3%; Pred. No. 9.4e-58;
 Matches 159; Conservative 69; Mismatches 105; Indels 43; Gaps 10;

QY 119 MNSMPKRLCVCGDIASGYHYGVASCEACKAFKRTTIOGNIIEYSCPATNECEITKRRK 178
 Db 176 MESAKETRYCAVCNDYASGYHYGVASCEGCAFFFRBSIQGHNDYMCPATNQCTIDKRRK 235
 QY 179 SCQACRFEMCKIKVMKEGVELDVRGGRQ-KYR-RIDAEN-----SPYLN 223
 Db 236 SCQACRLKCYEGVMKGGIRKDR-RGGRMLKHKRRQDDGGRNEAVPPGDMRSANLWPS 294
 QY 224 POLVOPAKK-----PYNKIVSHLVAPEKTYAMPDPVPSDIALTTLCDLADREL 276
 Db 295 PLTIKHTKNSPVLSLTDQWISALLREBPITISRYPTPLSGASMGILLTNADREL 354
 QY 277 VTIIGMAKHIPGSESTSLADQMSILOAMEIILGVYRSLSFEDELVYADYIMEDQ 336
 Db 355 VHMIMNAKRYVGFIDLSDHDOVHLECAWLRILMTGLVWRSVEHPGKLLFAPNLLDRNQ 414
 QY 337 SK-LAGLIDLNNAILOLVKKYKSKLEKEEVTYKATALANS-----DSMH 381
 Db 415 GKCYEGWEIFDMILATSSRRMMNLQGEFEVCKSIILNSGVYITLSITKSIEEKDH 474
 QY 382 IEDVEAVOKLODYLHEALQDYEAGQMEDP-RRAGKMLTLPRLRQSTKAVQHPFYNIK 440
 Db 475 IHRV-LDKITDTLIHMA--KAGLTLQQHRRILAQLLISHFRMSNKGMEHLYNNKC 530
 QY 441 EGKVMHKLFLBMLEA 456
 Db 531 KNVVPDYDLLEMLDA 546

Search completed: August 3, 2004, 12:43:43
 Job time : 46 secs